1 CCATTOCAAA CAAGTCAGGA AAGCCTGCAC AGGACTGGAT AAATAATTAA 51 GAACAGAGTG TTCTGAACAT CAACACAAAG TGGAACAACC TTAAGCTGAA 101 GGTACAGTAT ATTATTTACA CTGAAGGGCC TTGTGTGTGG ACAAGAAGC 151 GCTGACAGCT CAAATGGATC CCATGGAACT CAGAAATGTC AACATCGAAC 201 CAGATGATGA GAGCAGCAGT GGAGAAAMIG CICCAGATAG CTACATCAGG 251 ATAGGAAATT CAGAAAAGGC AGCAATGACC AGTCAATTTG CTAATGAAGA 301 CACTGAAAGT CAGAAATTCC TGACAAATGS ATTTTTGGGG AAAAAGAAGC 351 TGGCAGATTA TGCTGATGAA CACCATCCCG GAACCACTTC CTTTGGAATG 401 TOTTCATTIA ACCTGAGTAA IGCCATCATS GGCAGTGGGA ICCIGGGCTT 451 GTOCTATGCC ATGGCCTACA CAGGGGTCAT ACTITITATA ATCATGCTGC 501 TTGCTGTGGC AATATTATCA CTGTATTCAG TTCACCITTT ATTAAAAACA 551 GOCAAGGAAG GAGGGTCTTT GATTTATGAA AAATTAGGAG AAAAGGCATT 601 TGGATGGCCG GGAAAAATTG GAGCTTTTCT TTCCATTACA ATGCAGAACA 651 TTGGAGCAAT GTCAAGCTAC CTCTTTATCA TTAAATATGA ACTACCTGAA 701 GTAATCAGAG CATTCATGGG ACTTGAAGAA AATACTGGAG AATGGTACCT 751 CAATGGCAAC TACCTCATCA TATTTGTGTC TGTTGGGAATT ATTCTTCCAC 801 TTTCGCTCCT TAAAAATTTA GGTTATCTT3 G TTATACCAG TGGATTTTCT 851 CTTACCIGCA TGGTGTTTTT TGTTAGTGTG GTGATTTACA AGAAATTCCA 901 AATACCCTGC CCTCTACCTG TTTTGGATCA CAGTGTTGGA AATCTGTCAT 951 TOAACAACAC GCTTCCAATG CATGTGGTAA TOTTACCCAA CAACTCTGAG 1001 AGTTCTGATG TGAACTTCAT GATGGATTAC ACCCACCGCA ATCCTGCAGG 1051 GOTGGATGAG AACCAGGCCA AGGGCTCTCT TCATGACAGT GGAGTAGAAT 1101 ATGAAGCTCA TAGTGATGAC AAGTGTGAAC CUAAATACTT TGTATTCAAC 1151 TOOCGGACGG CCTATGCAAT TCCTATCCTA GTATTTGCTT TTGTATGCCA 1201 CCCTGAGGTC CTTCCCATCT ACAGTGAACT TAAAGATCGG TCCCGGAGAA 1251 AAATGCAAAC GGTGTCAAAT ATTTCCATCA CGGGGATGCT TGTCATGTAC 1301 CTGCTTGCCG CCCTCTTTGG TTACCTAACT TTCTATGGAG AAGTTGAAGA 1351 TGAATTACTT CATGCCTACA GCAAAGTGTA TACATTAGAC ATCCCTCTTC 1401 TONTOGETTOS COTOCONOTO CTTGTDDDA' TANCACARAC TGTGCCCAFT 1451 GTCCTCTTCC CAATTCGTAC ATCAGTGATC ACACTGTTAT TTCCCAAACG 1501 ACCOTTCAGO TGGATACGAC ATTTCCTGAT TGCAGCTGTG CTTATTGCAC 1551 TTAATAATGT TCTGGTCATC CTTGTGCCAA CTATAAAATA CATCTTCGGA 1601 TTCATAGGGG CTTCTTCTGC CACTATGCTG ATTTTTATTC TTCCAGCAGT 1651 TTTTTATCTT AAACTTGTCA AGAAAGAAAC TTTTAGGTCA CCCCAAAAGG 1701 TCGGGGCTTT AATTTTCCTT GTGGTTGGAA TATTCTTCAT GATTGGAAGC 1751 ATGGCACTCA TTATAATTGA CTGGATTTAT GATCCTCCAA ATTCCAAGCA 1801 TCACTAACAC AAGGAAAAAT AC

#### FEATURES:

5'UTR: 1-163 Start Codon: 164 Stop Codon: 1805 3'UTR: 1808

### HOMOLOGOUS PROTEINS:

Top BLAST Hits:		
	Score	E
CFA 145000033337444 /altid=gi 12017941 /def=gb AAG45335.1 AF295	975	0.0
CFA 114000033649823 /altid=gi 10945621 /def=gb AAG24618.1 AF293	5 ) 7	e-169
CFA[150000003782430 /altid=gi[8677401 /def=gb]AAF75589.2[AF1736	591	e-168
CFA 14300000002720069 /altid=gi 8248427 /def-gb AAF74195.1 AF2496	587	e-166
CFA[37::::0000006802 /altid=gi[7243:45 /def=dbj[BAA92620.1] (AB03	513	e-164
CFA[130.00005069215 /altid=gi[5870333 /def=ref[NP 006832.1] tran	500	e-140
CFA[830000001154721 /altid=gi 7406050 /def=gb[AAF61849.1[AF15985	496	e~139
CFA 65090019404613 /altid=gi 9505337 /def=ref NP_061849.1  amin	495	e-139
CFA 100:000004435450 /altid=gi 8926332 /def=gb AAF81797.1 AF2730	492	e-138
CFA 335001093689635 /altid=gi 11434147 /def=ref XP_006635.1  hy	430	e-134
EST:		
gi 10934204 /dataset=dbest /taxon=96	1072	0.0
gi 10286121 /dataset=dbest /taxon=96	718	0.0
gi[9872-34 /dataset=dbest /taxon=960	б <b>Н</b> ()	0.0
gi 2656674 /dataset=dbest /taxon=9606	549	e-154
gi[9882497 /dataset=dbest /taxon=960	541	e-151
gi n8th41 /dataset=dbest /taxon=9606 /	525	e-147

## EXPRESSION INFORMATION FOR MODULATORY USE:

library source:

Expression information from BLAST dbEST hits:

gi|10934204 Whole embryo (mainly head)

gi|10286121 Hepatocellular carcinoma

gi|9872634 Non-cancerous liver

 $g_{\pm 1} 2000014$  retal liver spleen

gi|9882497 Non cancerous liver

gi|689641 Liver

Expression information from PCR-based tissue screening panels:
Mixed tissue (Brain, Heart, Kidney, Lung, Spleen, Testis, Leukocyte)

```
MDPMELENVN IEPDDESSG ESAPDSYIRI GNSEKAAMSS QFANEDTESQ
51 KFLTNOFLGK KKLADYADEH HPGTTSFGMS SFNLSNAIMG SGILGLSYAM
101 AYTGVIFFI MLAVAILSL YSVHLLKTA KEGGSLIYEK LGEKAFGWPG
151 KIGAFVSTYM CNIGAMSSYL FIIKYELPEV IRAFMGLEEN TGEWYINGNY
201 LIIFVSVGII LPLSLKNLG YLGYTSGFSL TCMVFFVSVV IYKKFQIPCP
251 LPVLESVGN LSFNNTLPMH VVMLPNNSES SDVNFMMDYT HRNPAGLDEN
301 QAKGSLHDSG VEYEAHSDDK CEPKYFVFNS FTAYAIPILV FAFVCHPEVL
351 FIYSELFDRS BRKMQTVSNI SITCMLVMYL LAALFGYLTF YGEVEDELLH
4C1 AYSKVTTLDI PLLMVRLAVL VAVTQTVPIV LFPIETSVIT LLFPKRPFSW
451 IRHFLIAAVL IALNNVLVIL VPTIKYIFGF IGASSATMLI FILPAVFYLK
501 LVKKEIFFSP QEVGALIFLV VGIFFMIGSM ALIIIDWIYD PPNSKHH
```

### FEATURES:

## Functional domains and key regions:

[1] PI 0000001 PS00001 ASN\_GLYCOSYLATION N-glyc-sylation site

```
Number of matches: 5

1 83-86 NLSN
1 050-063 NLSF
2 64-267 NNTL
4 076-079 NNSE
5 469-372 NISI
```

[2] PDOC00044 PS00004 CAMP\_PHOSPHO\_SITE cAMP- and cOMP-dependent protein kinase phosphorylation site

100 500 KKET

[3] PDCC00005 PS00005 PKC\_PHOSPHO\_SITE Protein kinase C phosphorylation site

```
Number of matches: 7

1 33-35 SEK
2 49-51 SQK
3 129-131 TAK
4 290-292 THR
5 360-362 SRR
6 473-475 TIK
7 506-508 TFR
```

[4] PD0C00006 PS00006 CK2\_PHOSPHO\_SITE Casein kinase II phosphorylation site

```
Number of matches: 5
18-21 SSGE
2 22-25 SAPD
3 129-132 TAKE
4 305-308 SLHD
5 309-312 SGVE
```

# [5] PDOCO00008 PS00008 MYPISTYL N-myristoylation site

Number of matches: 6

1 95-100 GLSYAM
2 153-158 GAFVSI
3 164-169 GAMSSY
4 186-191 GLEENT
5 296-301 GLDENQ
6 482-437 GASSAT

[6] PDOC00009 PS00009 AMIDATION Amidation site

58-61 LGKK

Membrane	e spann	ing str	ructure	and domains
Helix	Begin	End	Score	Certainty
1	79	99	1.125	Certain
2	102	122	2.503	Certain
3	153	173	1.197	Certain
4	197	217	1.785	Certain
5	22.3	242	2.123	Certain
6	332	352	1.240	Certain
7	370	390	2.166	Certain
8	414	434	1.301	Certain
G	123	473	1.520	Cortain
10	476	496	2.166	Certain
11	515	535	2.628	Certain

## BLAST Alignment to Top Hit: >CRA|145000039337444 /altid=gi!12017341 $/def=gb|AAG45335.1|AF295535_1$ (AF295535) amino acid transport system A3 [Rattus norvegicus] /org=Rattus norvegicus /taxon=1)116 /dataset=nraa /length=547 Length = 547Score = 975 bits (2492), Expect = 0.0 Identities = 473/547 (873), Positives = 508/547 (923) MDPMELFHVNIEPDDESSCHESAPDSYIRIGNSEKAAMSSQFANEDTESQFFLTNGFLGK 60 MDP+ELF+VNIEP ++S S +S Y +GNSEK AM SQFANED ESQFFLTNGFLGK Sbjct: 1 MDPIELFSVNIEPYEDSOSVDSIQSCYTGMGNSEKGAMDSQFANEDAESQFFLTNGFLGK 60 Query: 61 KKLADYADEHHPGTTSF3MUSFNLSNAIMGSGILGLSYAMAYTGVILFIIMLLAVAILSL 120 F. L. DYADEHHPGTTSFGMUSFNLSNAIMGSGILGLSYAMA TO++LF+IMLL VAILSL Sbjct: 61 KTLTDYADEHHPGTTSFGMUSFMLSNAIMGSGILGLSYAMANTGIVLFVIMLLTVAILSL 120 Query: 121 YSVHLLLKTAKEGGSLIYEFLGFFAFGWPGKIGAFVSITMONIGAMSS/LFIIKYELPEV 180 YSVHLILKTAKEGGSLIYEKLGEKAFGWPGKIGAF+SITMQNIGAMSSKLFIIKVELPEV Sbjct: 131 YSVHLLLKTAKEGGSLTYEKLGEKAFGWPGKIGAFISITMONIGAMSSYLFIIKYELPEV 130 Query: 181 IFAFMGLEENTGEWYLNEGGYLTIFVSVGIILPLSLLKNLGYLGYTSGFSLTCMVFFVSVV 240 IF FMGLEENTGEWYLNGNYL++FVSVGIILPLSLLKNLGYLGYTSGFSLTCMVFFVSVV Sbjct: 131 IRVFMGLEENTGEWYLNGNYLVBFVSVGIILPLSLLKNLGYLGYTSGFSLTCMVFFVSVV 240 Query: 241 IYKKFQIPCPLPVLDHSVGNLSFNNTLPMHVVMLPNNSESSDVNFMMDYTHPNFAGLDEN 300 IYKKFQIPCPLPVLDH+ GNL+FNNTLPMHV+MLPNNSES+ +NFM+DYTHF+F GLDE Cbjet. 211 IYKKEÇIPCFLFVLDAN.AMLITANTLFMAVIMLFANALSIGMARMAD: 1460FLGLDEK 300 Query: 301 QAKGSLHDSGVEYEAHSDDKCEFKYFVFNSRTAYAIPILVFAFVCHPEVLPIYSELKDRS 360 A G LH SGVEYEAHS DKC+PKYFVFNSRTAYAIPIL FAFVCHPEVLPIYSELKDRS Sbjct: 301 PAAGPLHGSGVEYEAHSGLKCQPKYFVFNSRTAYAIPILAFAFVCHPEVLPIYSELKDRS 360 Query: 361 RRKMQTVSNISITGMLVMYLLAALFGYLTFYGEVEDELLHAYSFVYTLDIPLLMVRLAVL 4.10 RREMQTVSNISITGMLVMYLLAALFGYL+FYGEVEDELLHAYSEVYT D LLMVRLAVL Sbjct: 361 RRKMQTVSNISITGMLVMYLLAALFGYLSFYGEVEDELLHAYSMVYTFDTALLMVRLAVL 420 Query: 421 VAVTQTVPIVLFPIPTSVITLLFPKRPFSWIRHFLIAAVLIALNNVLVILVPTIKYIFGF 480 VAVT TVPIVLFPIRTSVITLLFP+RPFSW++HF IAA++IALNNVLVILVPTIKYIFGF Sbjct: 421 VAVTLTVPIVLFPIFTSVITLLFPRRPFSWVKHFGIAAIIIALNNVLVILVPTIKYIFGF 480 Query: 481 IGASSATMLIFILPAVFYLYLVKKETFRSPQKVGALIFLVVGIFFMIGSMALIIIDWIYD 540 IGASSATMLIFILPA FYLELVKKE RSPQK+GAL+FLV GI FM+GSMALIIIDWIY+ Sbjct: 481 IGASSATMLIFILPAAFYLKLVKKEPLRSPQKIGALVFLVTGIIFMMGSMALIIIDWIYN 540 Query: 541 PPNSKHH 547 PPN HH Sbjct: 541 PPNPDHH 547 >CRA|114000033649823 /altid=g1|10945621 /def=gb|AAG24618.1|AF298897\_1 (AF298897) amino acid transporter system A [Homo sapiens] /org=Homo sapiens /taxon=9606 /dataset=nraa /length=506 Length = 506Score = 597 bits (1522), Expect = e-169Identities = 315/549 (57%), Positives = 383/549 (69%), Gaps = 46/549 (8%) Query: 1 MDPMELRNVNIEPDDESSSGESAPD---SYIRIGNSEKAAMSSQFANELTESQKFLTNGF 57 M E+ +I PD++SSS S D SY +++AA+ S +A+ D E+O FL

MKKAEMGRFSISPDEDSSSYSSNSDFNYSY----PTKQAALKSHYADVDPENQNFLLESN 56

Query: 58 LGKKKLADYADEHHPGTTSFGMSSFNLSNAIMGSGILGLSYAMAYTGVILFIIMLLAVAI 117 LGERE Y E HPGTTSFGMS FNLSNAI+GSGILGLSYAMA TG+ LFII+L V+I Sbjct: 57 LGKKK---YETEFHPGTTSFGMGVFNLSNAIVGSGILGLSYAMANTGIALFIILLTFVSI 113 Query: 118 LSLYSVHLLLETAKECGSLIYEKLGEKAFGWPGKIGAFVSITMQNIGAMSSYLFIIKYEL 177 SLYSVHLLLKTA EGGSL+YE+LG KAFG GK+ A SITMONIGAMSSYLFI+KYEL Sbjct: 114 FSLYSVHLLLKTANEGGSLLYEQLGYKAFGLVGKLAASGSITMONIGAMSSYLFIVKYEL 173 Query: 178 PEVIPAFMGLEENTGEWYLNGHYLIIFVSVGIILPLSLLKNLGYLGYTSGFGLTCMVFFV 237 P VI+A +E+ TG WYLNGTYL++ VS+ +ILPLSL +NLGYLGYTSG SL CMVFF+ Sbjct: 174 PLVIQALTNIEDKTGLWYLNGNYLVLLVSLVVILPLSLFRNLGYLGYTSGLSLLCMVFFL 233 Query: 238 SVVIYKKFQIPCPLPVLDHS76TILSFINTLPMHVVMLPNNSESSDVNFMMDYTHRNPAGL 297 VV1 KKFQ+PCP- + 11 + 11 TL ++P Sbjct: 234 IVVICEKFQVPCPVEAA--LIMETINTTLTQPTALVP------ 269 Query: 298 DENQAFGSLHDSGVEYEAHSDDFCEPKYFVFNSRTAYAIPILVFAFVCHPEVLPIYSELK 357 + + + + + 10 C P YF+FNS+T YA+PIL+F+FVCHP VLPIY ELK Sbjct: 270 -----ALSHNVTENFGCRPHYFIFNSQTVYAVPILIFSFVCHPAVLPIYEELK 317 Query: 358 DRSPFKMQTVSNISITGMLVMYLLAALFGYLTFYGEVEDELLHAYSKVYTLDIPLLMVRL 417 DRSRF+M VS IS M +MYLLAALFGYLTFY VE ELLH YS + DI LL+VRL Sbjct: 318 DRSRRFMMNVSKISFFAMFLMYLLAALFGYLTFYEHVESELLHTYSSILGTDILLLIVRL 377 Query: 418 AVLVAVTQTVPIVLFPIFTSVITLLFPKPPFSWIRHFLIAAVLIALNNVLVILVPTIKYI 477 AVL+AVT TVP+V+FPIF+SV LL + FSW RH LI ++A N+LVI VPTI+ I Sbjct: 378 AVLMAVTLTVPVVIFPIFSSVTHLLCASKDFSWWRHSLITVSILAFTNLLVIFVPTIRDI 437 Query: 478 FGFIGASSATMLIFILPAVFYLKLVKKETFRSPQKVGALIFLVVGIFFMIGSMALIIIDW 537 FGFIGAS+A+MLIFILP+ FY+MLVKKE +S QK+GAL FL+ G+ M GSMALI++DW Sbjct: 438 FGFIGASAASMLIFILPSAFYIKLVKKEPMKSVQKIGALFFLLSGVLVMTGSMALIVLDW 497 Query: 538 IYDPPNSKH 546 +++ P H Sbjct: 498 VHNAPGGGH 506 Hmmer search results (Pfam):

	Description		E-value	
PF01490	Transmembrane amino acid transporter prote	in 187.0	2.9e-52	2
CE00398	E00398 CD53	4.0	4.8	

### Parsed for domains:

	or acmar							
Model	Domain	seq-f	seq-t	hmm-f	hmm-t		score	E-value
CE00398	1/1	90	110	 1	23	[.	4.0	4.8
PF01490	1/2	99	236	 1	179	[.	58.9	2.5e-14
PF01490	2/2	305	529	200	467	1	133 0	30-36

1 AGCTTAGCAA TATGGATCAA GAGGTCCAAT ACCTGATTAA TAAAAGTTTC 51 AGGAGTAAAC AAAGGYSANG PRATINGTITT TITAAATAGT AGAACTTTTT 101 TTATTTTAG AAAA DOT OOC TTOTATAGAA GAAAGACAAG COTTTTGATT 151 GGGCCGTCTG CATGCTSAFT AFGACGAATT TTAAAAGCGA CTCACATCTA 201 GTCACGTCGT GATGAAAHAA FAAGBATAAA AATTCTGAAA FCCTCAGAAA 251 ACCATEGATA AATTATOOAT AWABAAATAA GAGOCAGACT CATOAATAGA 301 AGCTAGAAGA GAGAAGTITC TTOAMTATTO TGAAGGAAAA TGOTTOTGAA 351 TCTAGAATTC AAACAATTAA CAAATTTGA AGGCAAAATA AAGAATTTTC 401 CAACATGAAG CAACTCAGAA ATTIMATTTA CAGACATAGG CTCATTGTGT 451 GAAAAAAGTT ATTCAAG HA THABHITAGC ATAATGCAAA ATAAACTGAA 501 GAAAGAAGAT AGAATGC ST TEAA BAAACT AGCAGCTGAG CAAGACTCAG 551 AGGTTGGAGG AGGAAGCCAT TOA MATGAG AAAGAGCATA GAAAATTTGC 601 TTTCAAAGTT TTGGTAATAT ADAADTATAT TTCACTTATT ATGTAGTCAA 651 ATACACCACT TTGTCTTTAG GECATACTAT TTATACAGTG ATAATACTGT 701 AATTGCTGCT TATTGGTTTT COAT TTTAG AAACAACCTA CAGGCAAGTT 751 ATGACACTTG TTTCACACAA CAA ANTGAAA ATATTATGAT TOTCAAATTG 301 TAAAAGTATT TTATTAACTA AAATWATTAG GAGTGTAGGA GAAGGAAGGA 851 AAGAAAGAAA AAGTATGUTA AUGUUDTTAT TITTITATGGG TAAGGAGTGT 901 AAAATCAGTA AACCAAGTAA AAAAAGCITT AGTGAATTAT TCAGATCTAG 951 AATGGCTAAC TTTAAGTAAC AAGGTAAAAAA CAGAAACCGT CAATAGTGGT 1001 TGCTGCTGGG AAGTGAGACT GUTACTGTGT GAAGAATGAG GAAAACCTTT 1051 GTACTCATTT AGTGAGTTTC ITTTTTTTTT CTTTTACCCA TATGCATGTC 1101 TTACTTCTAT TCTCTCTCAG CTTTTTAACCT GCTTCTTTTC ATCTTTTATG 1151 TATATACATT TAGGCTGCCT TATATTAATA ATAGTTTCAT TTTTGTTCCT 1201 CCTGCTTAAA ACACTGTGTG CTATTTTTTT AAATTCTGAG AACTGCTTTC 1351 TTTATTCTA GACAATTCTC TOCCATTATC TCTTTCTGTT TTGTCTCACC 1301 CTAGTCTCAC AATTCTCTAT ATTCCAATGA CTATCAGTGT ATATTTGAAC 1351 TTGTAATTCT TATTTTTTC CCATTCCTCT TAACTTCTTA TTTGTATTTT 1401 TOTTTTTTTA ATCTCTTONE GERMANET GAGEGATTEC CACAGATCTG 1451 TCTTTCAATT TTATAAGTCT TCCTTCAGCT GAGTTTTTTT AAATTTCAAT 1501 GATTCTATTT TTTTCTTTTT TTTAAGAATT CCTTTTTTTG ACTCTTTTTG 1551 CAACAGCCTG TTCTCCTTTT ATATTCCTTT ATAATGTTTT TATTCTGTGA 1601 AAGTTATTCT CTTATTTTGA ATGTTTTCTT TCAAAATGTC TTTCTTTTTA 1651 TTAATTTAAT GTAAAAGTCC CTTTTAAATT GCTTTGTTAT TTGTAGTTCC 1701 TTAGATGTGA ATTTTATCAT TTCTTGTGCC TACTGGCACT CTTGCTAGTG 1751 AGTTTCCATG TGTGTTCTAT ATGTTTTGTA ATTTGAGGAT GTGAACTTTT 1301 CTCAAGTGTG AGTTGCCTTT CAAAAAAGTA CTGCCATGGC ACTGGGTTGT 1351 GGAGGTATTC CCATGTGGTA GTTTCTGTTT GTCAGAGGAA TAGCACATTT 1901 TGTGACTTCT GGAGCAATTT TTATGTTAGT TTCTCTGCTC AAGATTTCCT 1951 TATCAAATGG GTATTGCACA TGTCATGACC ACACTTTTCA AGAATGATAG 2001 TGTTTCTCCT AATACGATGG TTCAACAATA ATTGAATGAA TCTAATGGTA 2051 AGAATTTCAG AAGAAATTAT ATCAACTACA TATAGTAGAT TCAAGGCATT 2101 TTTCAAAAAC ACAATGCCAG TCCACCCCTT TTCACTATAC AATTGAGGAA 2151 AATGAGGTCC CCAAATGTTA AATGACTTCT GCTGAGATCC AATGAATTAA 2001 AGGCAGAGCA GAGGCTAAAA TOTAGATOTO TTTGTTGTTA AAATACATTT 2251 TAATTTGACA CAGATGATGA GTAATGCTGA CCCAGAGGTA AATCTGAACT 2301 TTCTTTGTT ACTATTCTTA ACTTTGGCTT CAGGATCCAA GTGCCTAGAA 2351 AGTTACTTCC TAAACTTGAT CCTCACCTAT GTTGCATATT ATCAAGCATT 2401 TGGTGGTGTT AATTCTTTTA TGTCCAATTA AATTAAAGCA GTAATTTTCT 2451 TTCTAGTTAT TGCTAGTAGA GACACTGGTA GATTCTGCCT TGGTAGACCT 2501 TCCTCTGTCA ACAATTTACT TTTGTCTTCC TTTCTTTTAA AACATGTATC 2551 CCACTCACAA ATACCTAAAT TTCCTTGAAG ACTGCTGCCA TGTTTTAAGA 2601 TTTCTTTTT TTTCCATAGT GACTAGTAAA ACCTGCCATT TTCATTATAC 2651 ATAGGCACTC TATAAATATC TGCTAATTTA GCAATTATTA GTAATTTCCT 2701 TTCTTCTCTT CCATTTCTTC CTTTCTTGTA TTGGGTAAAG GAACATTTCA 1751 GGATTTGCTT ATGTAAAGIT TTCAGGAGTT TCTTTCCTTC CTCCCTTTTA 2801 CAGAGAGCAT ACAAAATGTA GATGATTCAT ATTCACTTAT TTCATTTAAA 2851 TAAAATTATA ATGATGTATG TTGTGTTCTG TTTGCAGAAC AGAGTGTTCT 2901 GAACATCAAC ACAAAGTGGA AGAAGCTTAA GCTGAAGGTA CAGTATATTA 2951 TTTACACTGA AGGGGCTTGI GIGTGGACAA GAAAGCGCTG ACAGCTCAAA 3001 TGGATCCCAT GGAACTGAGA AATGTCAACA TCGAACCAGA TGATGAGAGC 3051 AGCAGTGGAG AAAGTGCTOC AGATAGCTAC ATCGGGATAG GAAATTCAGA 3101 AAAGGCAGCA ATGAGCAGGT ATGGGGGTTAA AAATTACTAT GTTCCATGGA

3151 AAAATAAGAC AGGATGTGGA CATGGAAAAC AGGGTCTTGA TOOGAAGAAC 3201 TSBATTTATT ACAGGTAAAT TTOTGATAAC AATGATATTG ATOCTABCAC 3251 ATCAATTOOC TGGTCCTGAA ATACAGTGAT AATGTCAATC TCTTTTGTGA 301 CTSATTTAGA ATTGASGITA CAAIGICTTI GICTCCATTA ATAATGISTA 3351 ATAATTTTAA TTATTTTAGO CTAITGOTOO TOTTATOTTT CICAGATTOO 2401 FOTTTGAATG TIGGTACAGO TOOTSGITFO TGTAGGGATT CLITTICTOTO 3451 TAAAAGTATO OTOTGGGCAA GOTCACTCAC AACTACTATG COUTCACCCT 2501 CCAAATATAT GCCATATACC CAGCOPSTTA ASTITUTCTA CISAATITOA 3951 GATAATTATA TOTGAATSTO TAGDGCAOGT OTOTACTGGA CCATTAGDGT 0501 STCTAAATTG CCTCATTTAT AAASTYAAAC CYSTAATSTC TAAYACTSAA 3651 STOCTATOTT TOCCTOCAAA ASSINGSTOST SITCHASTAA TIGIDATIGIT 2/01 AGTGAAAATC ACTGCTATCA TGTAGCAACT CACTGAAAAG GCCCTAGGTG 3751 TAAACTTTGA COCACATAGO CAACGGTCAG TCATATCCAG TT GGTTTGAC 3801 CTTATTAATG CTTCAAATAC ACCTACTTTT CTGTACCCAT TOTACTGTGG 3851 TOTTACGTTA GGCCTACATT AAATGTGAGA CAGGGAGAGA GOCCTUATIT 3001 CTCTCCCTGT CTTACATTTT GOTCTCCTCT GTCTAGGCCT GTACAGTCCT 3951 GCAAGAGCAA TOTOTTACAA TITGOAAATTG AATCAATTTO CAICOTTAGA 4001 TAAAGCCCTT CTGCACCTCT CCAATAGCCA TAAGAGAAAG TAGACTAJAC 4051 ACACTGCTGG GCACGTAAGG TOCTTTGTGA TCTGTTCTTG ACCTGCCCCCT 4101 COTGTOOTGT TTTTTGOOGT STOCCTATTT GTTACTTGTT GOSTTSACTO 4151 ATTOTGOTOG AACTGOOTGG AATGAGTGAC CTGCTGOGGG TETCTGOGTG 4.01 TTGACACCTC TCATCCTTCA AGAATCAGCT CAACATCAGG TCTCCTATGC 4051 AGCCTTTCC AAATTACTCT ACTCCCCCAT GTAGAAGTGA CTGCCCCCCC 4301 TTCATGTACC CTCTCCCTGT GCAGATGTTA ATTACGCCAC TACTACAGGT 4351 TAATGGCCTC TGTGGTCCCA CCACCTGCCA CATTGTCTGG TCCATAGTGA 4401 GTGCACAATA GTTATTTGAT ANGTCAATTG ATTTCCCACA AAATGTTATA 4451 TCAAATTGTA CATGATTTAA GATGUTCAGA AGGGAATTTT TGAGCAAATC 4501 TAGGCGTGAA ATAGAGAATA TTGTGCTCAA ACAAAGACTT CTCATTTTAT 1151 TTADAACACC CAGGAAAATC CATCAGGAGA AACTACCGTT CTT.CTTCAA 4001 GTAGCTCAGT GCAATGAACT TTAGGGATGT CGGACTAGAG AGGCCACTGA 4651 GATGTAAATT ATAGCATTTT CTAAATTAGG TGACCCTTGA AGAAACACTA 4701 GGGTGCTAGA AGACAGGGCT TTGGAGTCTG CAGAGTAGTT GCCTGACTTT 4751 AGAGAAGCTG TTTGTCCTCT TTGAGCTTCA ATGGAAAATG TAAAATGGCA 4801 AACCAACAGC TGCTTTTCAA GGATGAGATG GGTGACCAGA ATATAGATGA 4851 CATTCAATAC TTTTTTATTA CTTCTCCTTC ACTGCATTAC CCTCAGTAAA 4901 TTGATTCAAA CCTGAGGATG TTTCTGAAAG GCATGCACAC AAATATGAGC 4951 TCTGCCGAGG TTGACAGAGT TAAAGGGGAC ACCCTCCTAA GAACTGTCAT 5001 AGTGTCATTC CACTTGATCC TCAAAAGCCA GAGTAGAAAG AGCATGAATG 5051 CTTTTCTTAA GCTTCATGCA ATGTGTTCCG AACCACTCAC AGTGACTTAC 5101 CTTTTATCTC CTGGCTTAAA CATAGGACAT CATTTTGCAG TTTTTAAAAT 5151 CAGTTTAAAG AGATGGGTTT TATCTATGTG TGGTTTGGAT TGAACCCTTA 5201 AATGTAAATT TTTGAGAAAT TCAACATAAT GTATTTATTT GTGATCATTA 5.251 TACTTGTGTT TTCAATACAT GCTGGGTTTG GTATCAAAAC ATTTAACATA 5301 CTGGGGACAT TTCTCATCTA TTTTATACAA TCTTGGCATG TTAAATGACT 5351 ACAACTCATC TCATGCCAAA ATAAGAACAT GCAAATGCCT CAAAGAAAGA 5401 AAATCTGTTT ACTTTCAAAT TCTCAATTTT AAAAACTACT ATGGAATACA 5451 GATTTTAGTT TATTGATTAA AATAAAGATT CCAGAGTTTA AATTCTAGGT 5501 GGCACTTTTG TTTTTATAGT CCTCAGGCCC ATTTTAGGCT TCATTTTATC 5551 CTGTCATCTC AGTCTCCAAC TGTGAACATT ATGTACCAGT CITCACATAG 5601 CAGGTACATT AATTACAGAC CATTAATGTA AACCACAAAA GAGTGGTGGG 5651 CAGTGGGTGG GGGGTGAATG GAAATGGAAA GAGGCAACAA CIGAGGGCAT 5701 TGTGCTTTCT GTGAGAAATA TGGGGAGAAG GCTAGGAAAT GTTCTTAACT 5751 TGTGTACTCA GAGCTATTTA TGCCTTGAGT TCTAGAAAAG CACATACAAC 5801 TTTGTGGTTT CGTGTGCTGT TTCTATCTAC ATCTCATACT GTTTTCTATT 5851 CTCAAAAAGT AACCCTGTCA TCCTCTTTCC TCTCCAGATT ATTTTCAGGA 5901 TTAGCTTCTG TTATAAAAAA TAGCTTGTAC AGATCTCCTA CAATAATTAT 5951 TTTCTATTT ATTTCTAAGG TTTATTTATT TATTTATTGA GACAGACAGA 6001 GTTTCACTCT TGTGGCCCAT GCTGGAGTGC AATGGTGCAA TCTCGGCTCA 6051 CTGCAACCTC TGCCTCCCAG GTTCAAGCGA TTCTCCTGCT TCAGCCTCCT 6101 GAGTAGCTGG GATTACAGGC GCCTGCCACC ACACTCGGCT AACTTTTTGT 6151 ATTTCTAGTA GAGACGAAGT TTCACCATGT TGGCCAGGCT GGTCTTGAAC 6201 TOOTGACCTC AAGTTATOCA COCACCTCAG CCTCCCAAAG TGCTGGGATT 6251 ACAGGCGTGA GCCACTGTGC CTGGCCTCTA GGATTATATT AATAGAACAA

6301 TOTTCAATTA TITTATOTTI CITTATOTTI CITTICATGI AGGAAATGIC 5351 CTAAAATTTT CAAAGCCTCA ATTTGAAAGC ACTTTTAAAA TOATACAFAG 6401 TOGAGCATTT TATATAAAAA CAACTAAAAA GTCTGTGACA TTTTGCAGTA 5451 TAAAAATGCA ATGGCAGCAG CAGGCCTTAT TAATTGAGCC TCCTSGAAAT 5301 GTGGGTGGTC CTAGGTGCGT AGCCTCAAAG GCCCTGGGTT GTAACTGCAG 6551 GASCIGACCA GCADAGCICI ATAACCAAGI IGTACATOUT CIAGCOTETT 5501 TOCAAGAAA OCABAATCAC AACGCTCTGT GGATAGTGAC ATCTTAAAGT 5551 TTTOTTTCCC TOCCAACTOT TTTGCCAGTT CATTGAATCG CTTTAATAAT 5701 TTOOTTAGTT TCATTCATTA TCTGTTAATA ATCCATUTAC ATTITUAGAG 6751 TAATTAAAAC ACATACGCAC ACACAGAAAC AACCAACACA ACACACAGACA 6801 ACCACTGAAT TACTTTCCAG TAAGAGATGT ATGTATAAAT GATTCTACCA 6851 AAAAAAAAA AAGAAAGAAA ATACCAGCTA CAGGGCCCTG CCTGTGACTG 001 CTTGATGCCA GGGGGAGAAT GGGGTCTCCC CCTGGGTATG GETERRATG 6951 GGCCTGCTGC TTCACCTTTC TGAGCCACAG TTCCCTATAG GUARACTITG 7001 AACATCAGAT GAGATAAGGA TCACAGTGCC TAGGCATTTA AFAAATATTC 7051 GTTGAATTAA TAAAATCATC TGATTATGGT ATGGTACTAG CCCAGAAAAT 7101 TCTGTCATAA CCCTGTACTC TTTCTTTGGA AGGGCTCTAA ATCGGAACAC 7151 AATTAGTTGT AGTOTOTTGC ATAGCTAATG TGAGAAAGAG GGAATOTGGT 7201 ATAAACAATT TTTTAACTAA AAATAATATT TOOTYOOTTI ATAACATOOT 7251 TOTTOCATOO CAAAGTATAG TTGTAAATGG AACTCAAAAT TGTTGGTCTG 7301 GAATGACCGT TAGTGTGAAG GAGGAAAAGA AAATTGGGGUT GTDTTATTTC 7351 CCCTCCTCTG ATTCAGTTAC TTAGATCACC TGAAACATAC ATATGATTCA 7401 GAGCATATAT TTAGATGTTT TCACTTTCTT ATTTGTGTGT GTGTGTGTTC 7451 AGTCAATTTG CTAATGAAGA CACTGAAAGT CAGAAATTCC TCACAAATGG 7501 ATTTTTGGGG AAAAAGAAGC TGGCAGATTA TGCTGATGAA CACCTAAGTG 7551 AATCTATGCT TTCAGGCAAT AAACGGGACT GAGGGTGTCT GATCTAGCTA 7601 GGTCTCTGTG GGAAAACAAT GTGACTGAAA TTTTCCAAGC CTTGATCAGC 7651 ACATTCTGTG TTTATTCAGG CTCTTACTGG AATAAGGGCT TGTTTTTTCC 7701 TETTCCCCAT ATCCCTOCAT CARTCATTA TGARACITAL G.G.TILGG 7751 GGGAAATCAT TCTAACCCAA AGGTAATCTA CAATCATACA TGTTTTCCCT 7801 TCTTTATGTG ACTCCCCTTG TAATTTGTAT TTTTACTGAG GCCTCTGCTG 7851 AAACCAAGCA CTGCATTCCG TTGAAAATTA CATGCTTTTA TTGATGTTGA 7901 GTAATGGCTT TACTCCTGTA ATGTTATCTT AGTCTTCAAT TTT5GACTGT 7951 AATCTGCAGA TAATGTGAGA ATAAGGATAA CCCCTAAAGG TATGCCCTTT 8001 GGCAAATGTT TGCTTATAAT ACATCCCTTC TTTTTCAAGC ATCCCGGAAC 8051 CACTTCCTTT GGAATGTCTT CATTTAACCT GAGTAATGGC ATCATGGGCA 8101 GTGGGATCCT GGGCTTGTCC TATGCCATGG CCAACACAGG GATCATACTT 8151 TTTATGTAAG TGAATGTATA TGTCTACATT TGGTGATGAA GTCCATGCAT 8201 ACCTGGTGGC TTTTTCAATT AACAATCTCA AGTTTGATCT TTGTGAACGT 8251 GAAGACTCAG AGGAGGCTAA TCATGGCACT TGGTCACCCA ACCATCCCTA 8301 ACCCAACGGC AGAAAGTGTA TGTGCTCAAT CAACCAAAGT GCTGGAGCAG 8351 CCTCGCCAGA AGAATTTTGT TATTCAGTAA ATACTTGAAA TAATTTGGTG 8401 TTTAGCAACC AAAAAGATCT TTCCCAGAAG CAAATCTGAT TTTATCTCAT 8451 TCTTAGGAAA GAAGCAACCA AGCCTAAGAG CCCTGCATGC CCTTGCCTAC 8501 CTTATGTCCC ATTCCCTGTA CCCCTGTGCG ACAGATACAC TGUGCACACAT 8551 AGCCTTCTCT CCATCCTATG AAGATGCCAC ATTCCCTCTC ACCATTGGAC 8601 CTTTGCACAT GGTCTTGGAA CCCTCTTCTC TTCCTTCTFC ATCTAGTTAA 8651 CTCCTCATAT GTCAGTTCAG TCTCACCTGA ATACTGCGCG COCTGATCTC 8701 CATGACTGGG GCAAATCACC TTATCATAAC ACTCACCACA ATTTTAATGT 8751 TTTAGTGCCA TTTGTCTGAT TCATTTGGTT AATATCTGTC CCTCTTGCTG 8801 GACTATAAGC TCTAGAAAGT TGAGCCCATG TCTGTTTTTA CTCACCAATG 8851 TCTCTACCTC CAAACCTAGA GCAGTGCCTG GTACAGGCAA TATTTGTTGA 8901 GTGACCAAAC CTTATTCCTA AACCTACGTA CTTTCACCAA ACTTGTTCAA 8951 ATGCTGCCTA AGGGTAGCAG CATCTGGTAG TTGACCTGTA GGGTGGATAC 9001 TGCACTGTCT ATGACAGACA ACAACAGACG TTTATGTGCA TCATGTACAG 9051 CCTGGCATTT TCCAGGATAT AGTTGGCAGC AGTGGAATTC TTCACAAGAA 9101 TAAAGTCTGA TGTTAGGCAC CACTGTGGAC ACAGATCCTA ATCCCAAATG 9151 CAACGCTAGA GAGTTAAATA ACTGTCTAAG AATGCAACAT TTATATCACA 9201 AATATGTGCT GTTTATGTTC TGAATATCAC ATATGATTAG TAATCACACA 9251 GCTATTTGAG GGCTAAGCAT CAGGACTATA AATATTTCTA TTGTGTTAGT 9301 GCTTTGATTG AACTCTTTTA TGTATAATAT TCTTCAGUTG AATGGGTTTT 9351 TATATCAACT TTACTTTAT ATAAGCCATG TTTTGAAATA AACTAGGATT 9401 TTAATAATCT GAATTTTAAT AGCTATGTAT GTAGTCATAT ATTTGTATGC

9451 TTTTGTAATG TGCTTACCTC TAAGACAAAA AAACCTGCCT TTCCTTATTA 9501 ATTATACATA CCATTAAAAT GAATTAGGAA GTTACAGATC ACTGATGAAT 9551 AGAAATAGGA AAAACTTOOC OCAATCOCAC AGTCATAGAT CATCTTCATG 9601 AGAGAAGAAY GUUCCACUUT TYAAAATGAG GGCCTCATUT TAGGCUTATA 9651 AACACTTAGC AGATGAATTT GGTCAGAACA ATTAAATCAC TAAACATCAT +701 GGGGTGTGTT TTGTGTGTCT AAGTAGCCCA GACTGGATTA AGCTTTCTCI 2051 CTTAATTTAT AGGAAGTGAG AGASTATTTT AAAGSTTTTA CTDTTAGTAF 4801 TTTCTGCCAG AGMAAGTACA TGTFTAGAAT ACAGGGAATG CTCATTATTT 4851 TTCCAGGGAA CAAAATTATA TAATCTGAAT TACATTATTC CTTAAAAACA ++01 GTTAAGTTCA TAAGGCATAT GGAAAAATAT AGGAATAAGT CATTGGTTAG 1951 ACAGTTOTGG CAMACATACT CTATGGAAAA TAAGAGTGCA ACATAGCTAC 10001 AGGGGTTATA ANATTTATAA TTCATGGTCC AAATGTACAT TTGTAGTATT 10051 GATTTCATTG GGAATTACCA AGGGATTAGA TCAATTCTGG GGAAAGTGTA 10101 TTTTTTAAAA AGAAAGAAAG ATAAAGATTT TTTTTCTGAA TTGGAGGTAA 10151 AAGGCAGCAT TOCTOSTOCA TTTATTACGT AGATGCTTCT AT MACATTC 10001 TTATTTTTGT GUTOCAAATC TTGGATTTGG AAAAATAGCA ATGGGTATAA 19.51 ACATAAAGAA ADOATACATG CATGTGGGGA TOCTAACACO AGAAATGACT 19301 CTGAATGCAA ANAAAAAAA AAAAAAAAA GGGAATTTTO CTGOOCCATC 19381 CTTAGCTTTC TOTGOTTTCT CTATTATATA TGCAACTGCO TGCOCCTCTA 10401 TOTTACAAAG TACTIT DETAA TOTAATGCAC AGGATCAGCA GTAATGCAGC 10451 TCAGACTGCA TOCTTTCGCC TTTGGATTCC TAGATTT DAG ATTAAGGTTT 10501 AGTCAGGCTA TIGAATAGCC CTTCAATTCT AAGTGCTGAT GTGAATATCA 10551 TGCAAATATG ATGTADATAT TCCCATGTGC TGAGTAAGTA GATGTAGCAT 10601 TTGCTAATGT TGCTATACAT TTAGCATCTA AGTTATGAAC CAGATTCTAC 10651 CACTGGGTAA CATTAAAAAA AAGTTAGGGA CTTCAGGTAT GTAAAATATA 10701 GOAAATTOTA TOTOTAOGAO TTTAAAGGGT ATGTGTAGAG TOTOTGAAAAG 10751 AATTTOTCAG OOTOOOOCAA ATCCACATAG TTTTGGAAAG OTGATGATTG 10801 AAAAGATTAA TGTGATGCTT TATTGTAACA TCTAACATAA TTACATTTTA 19851 TITATIGIAG AMACTITATI ACCIACION TOTICOCCIA SCAGAATCAT 10901 GOTGOTTGOT GTGGCAATAT TATCACTGTA TTCAGTTCAC CTTTTATTAA 10951 AAACAGCCAA GGAAGGAGGT ATGCTACCAC TTGAGTCGAA CAGATTCTAT 11001 TTTAATTOTO ATAAAAGAGT ATTTCAGTCT GTTGCTTCAT AACCTTAGGA 11051 TGATTATAGT CASTTTCACA TTTCATTTTC TTCTGAGCCC AGTGACACGA 11101 TCTCTCAGTG TTTATAGTTG TTTGGGCAAG TGAGAGGCAG GAGTGAAAGT 11151 CAACTGGCTC AGGTTCAAGA CAAATAGAAA AAAGAAATTT CTGATATATG 11201 ATAGAAATAA CIGTITIGAC TIGCTACATG CAGCTAAAAT AAATAAAACC 11251 ATTGATTCTT GTTTGGAGAA CATTTTGATA TATTGCTTAT TGGTTTTTGA 11301 GGTTGCATCT TTTGGGCTTA TAATTTCTAT ATGATGTTTA TTTACATGTT 11351 TGAGACTOCA GUATGGAATT ATATGACAAA AATATTTTAG TOATTAAAAC 11401 AATCTCTTTA ACAAGGCTAT TTTATCTTTG ATTGTAGGGT CTTTGATTTA 11451 TGAAAAATTA GGAGAAAAGG CATTTGGATG GCCGGGAAAA ATTGGAGCTT 11501 TTGTTTCCAT TACAATGCAG AACATTGGAG GTAAGGGGAT ATACTTTCCA 11551 ATGGATOCCA TAAACTTTCT ATAGCGTGTT CAATAAATAA GAAAACTTAT 11601 GGCAATAAAC AGGCACTTTA GATACAGAAA AATTGCTACT TATAGTTCTT 11651 AAATTTTAAA ATGATAGTTT CTTAAATAGG TTTGTGTCCT GCTTTAATTA 11701 AAAACAGGAA TATOTAAGAA TGAAATAACA TATAAAAGGG TGCCAATTGA 11751 ATTCTAGAAT TAAAATATAA AATAAAAGCT TTCTTGATTT TTAATGTTAT 11801 TATAGCATGA ATTATTACTO TTAAAAATTG AAGAATTTGT GCTTATATCT 11851 GTCATTGACA AAACAGTTGA CGTTTTCTAT GTGTGACTGA GTTCGATTTA 11901 CTAAACTGAA AAGTGGGTGT CTGGGGGAAC ATAGCCAAAT GCTGTGGTCC 11951 TTGAAACGCA GOOTGCACTG AGCCAGCCCA CTAGACAGTG TOTOTGGAAG 12001 TTTACTAAGG CAAAAGTCTG GCTAGGCATC AAATGCACTA TAAACCCCGG 13051 TTTGTTGATT CTATGGATTC TTATAATTCC CACTGAATTA TCATTTCCAG 12101 TGTAGGACCT AGAAATATAT ATATATTT TTAACAATGT TCTCTCGTTG 12151 GTGTGTTTGG GGAGGAGGTT CATACTGTTT CTGTTGTGTG TTTGGGCCCTC 10201 AGAAGGCATO CAAACCCATA TITCAGATGT CCTGCCGGCT GCTTCCTGGC 12251 ACATGGCCCC AGOCATOTOC CCACATAATG ACACTTACTC COFCACCTCC 10301 FACCCAGTOC CHAAACCTGC TATTCTATTT CTCTGAFCTT TCTTTTCTCA 12351 GTGAATACCA CCAGCAGTCA TCCAGTTTCT GAGGGCAGAA ATCTGGATGT 10401 CAGOGTAAAT GITTGGTTTT CCCCAACTCT GCATGTCCAA TCAAATGGCA 12451 AAGTOTGTTO ATPIGATOTO TTACTTATOT CTTGAAOOFO TOOTGTGTGT 12501 COGTOCTCAT GACCACAGAT GATCACCATT TATAGCTCAG ACTATTGCAG 12551 TAGTOTTOTA AUTOSTOTTO CTGGCTTGAG TTTCCCCTGC TOTCAGATAA

13601 ACTOTAATTT GTTCTCCAGA TAAACTTTCT CAAATTTGAG TCTGTTTCTA 1.1651 CTTTTGTCGT GCATAAAATT CTTCAGCAIG CCTTTATTAT TTTCAAGGAA 10701 AAACTTAAAC TCATTGCACT GACACAGAT CTTCGTCTAG TTCTTCTGCT 12751 CAATCTTTCT AAACTTTCCT AGCAATGCCC ATATCTATCT ATCTTTATCT 10301 ATCTATCTAT CTATCTAT OF AUGUSTATIONAL STATISTATCT AUGATOMATIC 10351 AATTTATOGA TOATOTATAD COTACATSTO CTGTGTCAAA CCATAACAAA 10001 TTATATTTAT TOCCCTAACA GTACTATTTT AATATTTTTA AAAATCATCC 10951 ATGCCTTCTT TTCACAGGCT ACTTTCTCCC CTTGACTGTC TCTCAAAGTC 13051 CACACACATT TTCTCTCTCA CTCTGCTCAC CTGGTCTATT GCTCCTCTAG 13101 ACTGGTAAAT ACTAGTTOOT CTGGGCTCTC ATGGTCCTGT TTGTATCTAG 13151 TATGTTACTG TTTTCTAAAG GATATTTTAA AACACTTGAG TAGAGAATAA 13201 GCTTTTGGAG TCTGATGGAC CTGAATTTGA GTCTGTTTCT GTCACTATCT 13.251 GTGAACTTGG GAAGATCACT GTACTCCTTT GTCTGATTTT TTCATGTATA 13301 AAAATTACCT TACAAAGGCT ATTGTGAGGA TGAAATAAGG TAACATATGG 13351 CACATAATAA GTGTTCTGTA TATGCTTCTC TCCTCCCTGG TTCTCTGCTT 12401 CCATATCCAT GTCTCTGGAG TTGCCTGAAT TATTTTTTAA ATAGGCATTT 13451 AAAAAATTAT AAAACAAATA TATGATGATT GTGAAAAACT AAAACACTGC 13501 ATAAATATA AAATTACCAA GAAAAGTTTA TGTCAGTCAT CCTCAGAAAT 13551 AACTACTCAT AGGTTTTCCC CTATGCCTAA TTCAACAAAT ACATTGAATA 13601 TTGTTAGTAT TGGATCATCT TATGATACCG ATTTTCAGCT TTCTTTTTAA 13651 ATTTAACAAT ATGCCTTGAA TATATTTGCA TGTTATTCTT TTTAATGATT 13/01 TTTGAGGTTT CCATTACACA AATGTGCCAT AATTTGTTTA CAGTATCCTT 13751 ATTGATGAAC AGTTGGATTG TTTCTAATTT TTCACTGTTA TAAAAATGCT 13801 ACAGTAAATA CACTTGCACA GAGATCTTGC AAACAGGCAA CCCATTTTAA 13851 TAAATAAATT CACTGGAGTT ATCAAGGATT TCTGGAATGC AGAAATTTCT 13901 TTAGTAATCT ATCTAACTAT ACTCACCCTG ATAATGGATA GTTGGTAAGC 13951 AGATAAGTAA AATTCAGCCA TATCTTATGA TTTGTGTTAA AAAAATTTTT 14001 ATATGTTAAG ACTACAATCT TOOGTAGAAT TYGACAGTAA YATCAAAATT 14051 GTCTCATTCA TTTTACTGGT TTGGAGCCAT ATGCATATTA GCCCCCCAAA 14101 TCCCAACAAA TAGACCACTT TACATTTGTT TCAAACTCTC AGCUTTATCA 14151 AGGTTTAAAG TATCGAGCAT TTCATAGGAT TGCCTTATAG TTGGTCTAAT 14201 TTAACAACTG AAATAACCAG GCATAAGCAT AATTAACCCT GGACTCAAGA 14251 AGTTGAGTGG CAGCACCTCA GCTGTGGTTC AAAGCATAGC CACTACTACG 14301 CTTCTAAACA ATGGAATAAA GTATAAAGCG GTCTCTCAGT CAAGCCTCAC 14351 ACAGGTAAGA GGCGTGACTT TAAGGGAGTA AGATGAAATA TCGTAACATC 14401 ACCCCAGAAA TAATGCTCTC ACTTTGGTTA CTTTATTTGA TTAGTTGATA 14451 TTTGGCATAA GAGAAATCAC TTGTATTTCT CTATTTAACA ACTCTACATT 14501 TAGAACACTT AATTTTCTCA ATCCCCTAAA AAATTAACAT TTACTGCAGA 14551 TGTTTTCACA TTAACAGATT AATGTCTGGA TCATTCTGAA TTTTTGAAGA 14601 CCAAACATGT TAACATCACT GACATCACTG AAAACCAGCA ATTAATAGCT 14651 GTAACATTGA ATGGTACCTC ACCAAGCCAG CTAATCAGAA ATATCTCCTG 14701 TGTTCACACT CTGTAAGATT TAGCTTTAGC CAAGGTCTTT GCAAAGATTA 14751 ACCAAATAAT GTGTACAGAA GGTACATCCG CTATTGTAAA AATCATTTCA 14801 CTTTGACAGT ACAGAAGAAG CACCAGCCCT TCTGTTTTAG ATGTAGTCCG 14851 TCCTTTTCAA GCTGTATGAT TGTGGACATG TCAACTTAAC ATCTCGGAGT 14901 TTTTATATCT TCATCAGTGG AATGAGAATA ACAACATATA TCTTGTCATC 14951 TCACAGGGTT TTTCAGATGA TCAAATGAAG TAATGTGCAG AACTAACCAA 15001 TGTGGGGAAT TATTATCATC ACTGTTACTT TCATATGAAG TGAAGAAAAT 15051 ATTTTTAAAC TCAGTAGTTT AATTTACAAT TTAAGTATGT GTTTTAAAGT 15101 GCCTGTTAGC AAAAATTCAC TAGAAGGATG TAGGACACAC TTAAAGTTTT 15151 CATGTAAAAT TTGTGAGTTC TATTTTTAAC TGAATCTTTT GGCCATGTGT 15201 CAACAAATTA ACGTTATCCT TCACCAAATG GGTGGGCTTG AAAAAGGCGT 15251 GATGCATAAA TATTTACAGT TGTAGGCAAA ATTGTAATGT TATGTATATG 15301 AATACATATT CATTTTTCA GGGAGAAGGC TTGTAGATTT CATCAAGAAA 15351 TCTTTCACAA GAGTAGATAA TCATTCATGT ATCACTTACC TAGATGCTCA 15401 TGAAATTTTG CCACTTTATA TAATTCCTTA GTTAGCCAAA AGGAGAGTAA 15451 GATGAAGAGG GGGGAAAAAA AAAACTTCTT TGACAAAGAT GGAGAGAAGC 15501 TGTCATCTCT TGTATTCTTT TATCAATCCA GGAAGCCTTT GGTTTTGACA 15551 ATAAGTGGTC TGAGACTTTG TGTACTCCTC AGATAGGTCC CGGAGGACTA 15601 GATTGGTGCC CATCTGCAGA AAACCAGAGG GGATATATTG ACTCTGCAGA 15651 TCTGCCCTTT GATTCTGCCA TCTCTCAGCT GGCCCATGCC TTTTGTTGCC 15701 AGACTACTGC CCAAGTTATA GACACTAACA CAGGCACACT GAGTATGGGC

15751 TATGTTGATT TATAACTAAT GAGGGCAGAA CCTTAGAACT GCAGCTTCAC 15801 TGTAAACTTT GGAGCAGGAT TTAACACAGA ATCAGCCCTG ATACIGTTAA 15851 CAAAGGTCCA CCTGAAAGAG CTGGAAGGTC AAATGTCTAT CTTGGAAGAG 15901 AACTTGGAAG CAGTGCCAAA TACACAATGA CTTTTTTTTC CATTTTSGGGG 15951 APTAGATGTT CATCTTACAT ATCCCAAATG TCATAACTIG CTTGCALGTG 16001 ACTICAGIAC IGICCACACO ATTAAGCIGI CACATITICO ATTITAGCAA 16051 TGT-CAAGCTA CCTCTTTATC ATTAAATATG AACTACCTGA AGTAATCAGA 1-101 GCATTCATGG GACTTGAAGA AAATACTGGG TATGTCTTAT GCTCCCTCTG 16161 TGACATCAAG TGACTCATTC TACTTGGTCT TTTCTGATTC TAATATCCCT 10201 STOTOTOACT TOTAGAGAAT GGTACCTCAA TGGCAACTAC CTCATCATAT 16.351 TEGEGEGET ESGAATTATE CETECACTEE CGCTCCTTAA AAATTEAGGE 16301 AAAGATATTT TOTAACTGGA AATATTTTTA TTTTTATTTC ACATTTAAAT 16351 AGGTTAGCTA ATTGTAGATG COATATTCAC CTTCCAAAAT GCTTCTTCTA 16401 ACTICTAGGT TATCTTGGCT ATACCAGTGG ATTITCTCTT ACCTGCATGG 18451 TGTTTTTGT TAGTGTGGTA AGTGATGTGA TGACATGATC CTTGCAGGTT  $16501 \ \ \mathsf{GGTTAGCATG} \ \ \mathsf{AGTTTTTTG} \ \ \mathsf{TGCCTAAATT} \ \ \mathsf{AGTGTCCTCA} \ \ \mathsf{TTTTGTTCAA}$ 16551 GCACTICACT AATATGAAAT AGTTCTTGTA TCACAAGTGA TTTTCTTGTA 14501 GACTAATTTA GAGCAAAAAA AGAGCAGCTA CGATTTAAAG ATAGTTGAGG 15651 TAGAATATCA AAGCTACTAC TAATGGTTTG GTCTAGGCAC ACTGGTTATA 16701 TATGGGGAAA AAAGGAAAAC TICAAGCAGG AACATGACAA TAATCTGGCA 16751 TTTAGAACAG CAGAGGAGAG TCCCAGATGA GAAACAAGAA GGCTATATCC 16801 ATATTCACAT GAATCAGCCA TTCTCTCTTA CACATTCCAC CCATTAAGAG 16851 AGGACAAGAA CAGTGGGATT AAAGAAGAAA TCCTCCTCTC TAGGCCCCCTG 16901 ACAAAAGAG GAATTTCTTG CACTATCATG AATGCCAAAA TTTATAAAGG 16951 ATTTCCCCAA AGAGGTAAAG GAGAAGGAAA AAAAGTTTTG AAGACCCATG 17001 TCACCTTAGT TTGAAGAAAT AAGGAAATGA TCATCTTTCT CATGGAAGGG 17051 CATGAAAGAG GGTGGGAAGG ATTCTTGCAA AATATTGTCC TGTTAACTCT 17101 AAGAGGCAGG GCTGCCAATC ACAGCTCCAA CTCTTCCCTT AGAACAGAGG 17151 CTACACCAAC TTTACTTTCT CCATTACTCT AAAAGGAATC CCTAACTGAG 17201 TTCCCTCACC CCCCACCCTA TAAGCCACAC ATATGGATTC TTATTTCATT 17251 GTTTTTCTC AAAAAGCTGA TTTTTTTTC TTTTTTAATG ACTGAGTCTA 17301 GGTGATTTAC AAGAAATTCC AAATACCCTG CCCTCTACCT GTTTTGGATC 17351 ACAGTGTTGG AAATCTGTCA TTCAACAACA CGCTTCCAAT GCATGTGGTA 17401 ATGTTACCCA ACAACTCTGA GAGTTCTGAT GTGAACTTCA TGATGGATTA 17451 CACCCACCGC AATCCTGCAG GGCTGGATGA GAACCAGGCC AAGGGCTCTC 17501 TTCATGACAG TGGAGTAGAA TATGAAGCTC ATAGTGATGA CAAGTGTGAA 17551 CCCAAATACT TTGTATTCAA CTCCCGGGTA AGTGAGCGGT CCGGGCTTCT 17601 AATGAGTACA GTTATGTGTT TTCTAAGTTT TTATTCAATA AACTGAGATG 17651 GCCTGAGATC ACCATCTATG TTGGAATGCT AAACACGTGG TGTTGTCTTT 17701 GTTTTTCAGA CGGCCTATGC AATTCCTATC CTAGTATTTG CTTTTGTATG 17751 CCACCCTGAG GTCCTTCCCA TCTACAGTGA ACTTAAAGAG TAAGGCAGCC 17801 ATCATTTAG CATTCTAATT TGCTTTGAAA TTCTGCTCAT ATGTTCAAAG 17851 ATTCTTTAAC AGGAAACACA GTTTATAGCT TCCTCTTCAG AGAAAATATG 17901 TACTCCATCC ACTCCTCAGT AACATGCTTT AATCAGAAAG GTGGGAATCA 17951 GCCCACCACA GCACTACCTT ATCTTCTTC TCTCCTTTCT CTCCACCATA 18001 ATGGTTCAGG GGAGGGGTTC ATGGCAGGTG GACAAGGAGT CGATGGTTGT 18051 AATAATTTTG GCAGGTGTTG GGAATTTAAA TTTGAATTTT GTTCGGAAGA 18101 AATGATGTCA GCTGGACTAG AAATGAAAAC ACCCATGACG ACCAAAACTT 18151 ATGGTTAGGG GCAGCCTCGA TAAGCCAGTG ATGTCATTTA TAGTCAGCAC 18201 CTAACCCTTG TCTAGAACAC ATTCATTACA AGAGATGTGT CAATATCTGT 18.51 CCTTTGTTGT CTTATTTGTA CAATAGAGTC ACTGGCTAGA AAATCTTGTT 18301 TCTTCCAGCT GATGGTCTAT GGTTCATTTG TATTCTTTTC CCTTTGAAGT 18351 TGTTGATATT TGCTTGGGAA CAAAGGATAT GAACTCATTA TAGCTGTTTT 18401 CCTCTTTCCT TTAAGGGAGG ATATTATATA ATAATTCTCA ACTTCTTTAA 18451 TCTAGACATC AGTAACCTCA GTCTTCATTC TCACTAAATA GCAAAACTTT 18501 CCCCATAAAT TCTGATTTAC CTCATAAAAA ATTTCAGAAC ACTTTCAAGT 18551 ATTTTGATGT CTTTGATTTA CTTTGAAAAT TACATGTAGC AGTTACTCCA 18601 GAAGCCTGAC AATTGATCTT TGGCAGCCAG GTTCCTTCTA GAATGGTTTT 18651 CAGAAGCTTT TCAGGTAGTC TGGACTCCTG GCAGTAGTAC TTTGCTGACT 18701 CTACTAGGTT CTTTTCCTCA TTTAAAGTCA TCTCATTATG AAATGCAAAA 18751 GCTTTCTATG TTAGGAGCCT GTTTCATCTT TATGTTAATT ATATTCTTAT 18801 TCAGTGGGCA AGCTTACTGA CCTACGTGAA ATAGACTGTT CCTCTTCTAG 18851 GGAAATGATT GTTTTTAAGA CTGAAGGACT AGTGTTTAAG AAAAATGGAA

18901 ATGAATCCTC ATTAGCTCTC TAAGACAAAT TTAAATCAGC TATAAGTTTA 18951 TGTACTAAAT ATGTCTTCAT UATTAGCAAT ATAGATATAC TTTTTTATTA 13001 TRATTTCAT TTTGAAAAGT GATTTTTTTT TGTAAGTTTA AAAAACAAAG 19051 CTTGGTGTTC TTTCTTTTC TAGTCGGTCC CGGAGAAAAA TGCAAACGGT 19101 GTCAAATATT TOCATCACGG GAARGEITGT CATGTACCTG CTTGCCGCCC 19151 TOTTTGGTTA COTAACCTTO TEATGGTAGGT CACTOTGAAA GTCATTCTCT 19201 ATATGCAAAT COTTGTTAGG CTGGTCCTTG ACCTGGGTAG GTATGATTTT 19251 TAAAAATTGC CTTCTATAAG CATGCTCTAT AGATGACACA TATTCAATTA 19301 ATATACTATI TTAGTTTTGI CACTTGACCT GAGGAAATGG GGCCTGATTC 19351 AGCCTGGCTA ACAAGTTA A AGAACTTGTG AATTAACACC TATTTTATAA 19401 AAAATATCCC TCAAACAAAA TTATTTCCT CTAGGGATAG ATGATATTTC 19451 TCTGGCTAGA CTCCATAGTO CAACTCAGGC TACAAGTGAT GAGAATGAAT 19501 CCACTTGCAT GTGATAAAGC TOCTTTGATG GAATTATTAA CTGCCACACA 19551 AATAGCAGGG AAACTGOOAG GUCCTCAAGT TTGAATTTGC CTCCTCTTA 19601 CCAGTCAAGT CAAATCTGGG AGCTTGGGGAC TTTAGGTAAA ATTTCTGACA 19651 TATCCCATTC TATTTUTUA TACTAAATGA TTTCCTAAGA AAGAGGACAT 19701 GACAGAATTT CCTTCAATCT AAGAATGCAC CACCAAAAAA AAGTGACTAT 19751 GGCCACATTA GATTATGCCT GUAACATTTC CTCTCTGGCA TCTTAACAGT 19801 TCACAAAGGG AGTAGGATTG TACTCCTTCC ATGAAGTGTG GCCACATAAA 19851 CAGATTTCAT GGAATCACAT ATTGACCTGG TAGCATATGT TTACATGAAT 20001 AAATTGTGGA GGAAAAGATT CAGGAGTAGG CCACCATTTG CTTAGGTTTT 20051 TTTTCTATTT GACTAATATT TGACTATTAA CCAAACATGT GCTTTAGATT 20101 GGGCATTAAC TTTTTGCCGG TTGTGAAATA ATGAATGACG AGGTCAATAC 20151 TACTGAAGGT ATTTTCACTA CTTTTTGTCT GATCTTGAGG TGAAAATCCA 20201 ACTACGCTTG ATTCCATAGA TATTTTCTTG TTATTTGTGC TTGGAGTCCT 20251 GAATGAAGGT GTTTTCAAGT AGGGCTGCAT CTTCGTCTTA GAGTAGTACC 20301 CACTGGGAGA CCATCTAAAA ATTATAGTAA TTTATAGCCTG CACCTTAGTT 20351 ATACTTATTT TAATGAGTTT CATAAGACAA GCAAAAACTT GAAAGAGCCC 20401 AAAAATATCT GTTTTAGTGT GGTGATGGAG TCATAGTTGT TGAGCTTGAA 20451 AAAATGGTAG CAATCATTCA TOOTAGAGTT TACACACTGG GTTTGTAACC 20501 TGCATCAGGA GTGGCTGCAC AGGTAGGGAC AGGGGAGGTG GTAGGCTGGG 20551 AGAGACAATA TGTGGGGCTT GGGTCTCTCA TCCCCTTCAA CAAGAGCACC 20601 TTGGTCTCTG TCTGATTTGT AATTGCTTCT GTACAGCGGA GATAGATTTA 20651 TCACAATGTA AATGAGCTTG AGAGGCTCTT TATTTTGTAT TATACCTTCT 20701 GCAACGTTAT CAGCTTCAGG ACCTCTTTGT TCATTTGAAT GAAGGTTGCA 20751 TAGCTAATGA GCTCAGAGGC AAGACCAGAG GTGCCTGGAT TCCCAGGCCT 20801 AGGTCTTTTC CTCTGTTCTG TGTTCTCTCT ATAAAATGTT GCCATAAGTG 20851 ACCTGTGCTG ATTTGACAAC ACCAAGCGGT TTCATTCTCT TTTTCCTGTT 20901 GTAGGAGAAG TTGAAGATGA ATTACTTCAT GCCTACAGCA AAGTGTATAC 20951 ATTAGACATC CCTCTTCTCA TGGTTCGCCT GGCAGTCCTT GTGGCAGTAA 21001 CACTAACTGT GCCCATTGTC CTCTTCCCAG TAAGTACATA AGACTTTGAT 21051 GAAAGAAACC TACTTGACCC CATAAATTAG TACATGTGTT CTACCTTCAT 21101 TTTGATTTAA TTATAGGGTG AGTTTGCAAT TGCAATGCCT GAGGATATTA 21151 TTTTCCTATA GCATTTTGAG TCACTTAAAA TTGGCCATTT AATGTGTAGA 21201 TAGAGCAAGT AGTTTCAGGT GGTATTTTTA TAGTGTAGGA AAAAAATCAT 21251 AAAACTTATT TTTAAACTCA AAGTTGAAAA GTGGAGCTGG AGCTTCTGTC 21301 TTGTGGATTA GTAAAACTBA GTAGGAGTTC ATATAACTTT GGAACCTTGA 21351 AAGCCAAAAC CATATTAAGT TTCAAATCTT ATTAAATTTC ATCACAGTTT 21401 TGAAGGCATT TCATTTTTT TCCAGTTTGT TGTGCTGCAA TAATATACAA 21451 AAGTTGCCTT TTTTAACCTG ATGCCTTGAA GGCTAATGAA AAGGGGATTC 21501 ATGTTAAGTA AATTATATAC CAGAAAAAAA TTTTTCAAAA AACAGTTATG 21551 CTATCTATCA CATATCTCTC TCACACATGG CCTCTGCCAG ACTCACACCA 21601 GGTCACCCCT CCCTGGCATT TGTCATTGGT GTCAGTTTGT TCTGAGATCC 21651 CAGAGCAGAG CTGGTAGTGA AGATTTGGGC TGTGTGAGTT AAAACCACCA 21701 CCTAAGGATA AACACAGGTO TTCACCCTCC TGCCAGCTCC TGTTTCATAA 21751 ACACTGAATT TACTCATTCA TTIGAGGGGG AAAAAAATAA GTGACACAGT 11801 AACCAGCACT GTCCTGGACA TAATGTTCCA TACAGGGCTG GCATATGAAG 11851 ACTATTTCTA TAATGACACT GUGGTCACTT TAAATGCAGC TUGTGUGGUG 21901 AAATATATTT TGGCACATTC CTTTTTCATG AGTGCATGAA ATCAGATCCG 21951 TACTACTATG GTGGCTAATA TTTTACTCTT AAATCATGTC TTGCCTCTAA 22001 TATATCTGAA AGTATTTCAG ATGACATACA CATAGCTTTA GCCTAAAATC

20051 AGCTCCGTCT TGGGTACAAG ACAGAAGACA ACTATAAACA GAAGGTATAC 2.1101 GATAGGGTAA AATTGCCAGG CAAACAACTT CACTGAGAAA AGGATATOTG 22151 GAGCCCTTCT TTTTATGTCT AAAAAAATCA CTCACTAAAT TTTDGCACAG 20201 TGTAAGCATT CACATCATTS TAGAATCAAA GCATAAGAAA TCTGTGATGT 2.251 GOTTOTGTAT TGCTTTATTO ATATTCATAT AGTGTTTTCA AGCCATGGTT 22301 Traagggatt godagaares godatogtca cacagacago tgotaacagt 200351 TOAACTAGTG CAGCTCATAG COCAACACTG AGGGCTGCAA TTATTGTCAT 111401 GGGAAGTAAA AGTCATITAC TBATGAACAT TTCACCTCAG CATGGAAAAT .L.451 CCAAATCTCC CCTTAGAAAT TOTTACCCTA TGTGAGAAAT AAAGCACTGA 2.501 TATAAATCTG ACCATCAGGA ACAGCAATAG TGTGTAAACA ITAGATGCCA 22551 TTAGAACCAA AATTGACCAT AAGAACCAGA GTTCAGAAAA ATGACTAACT 22601 GCTGTCCTTC ATTATGTATT TOCACTCAAC ATTAGCATTT ATGAAACATT EP651 TTGCACATTA TCCTGTUCTO AUGCTTGCAA TGTTACATTT ATATAATCTG 20701 TGTAAGTGCT CCACTGCOCC ACAGAGTCAT AAGTCCCTGG GACTTGGTGA 22751 TGTGCACAGT GACTGGCACA GAGGGTGAGC TCTGTCGTGC TTGGGAAGAA 22801 AAATGGTCTT CAAATGAATC FTGCCTTGTC TTGAAATGTA TAAACTGCCT 22851 TTTCTAGCAA AAGCATAGAC AGTCTTTCCC TTGGTGACAT GTGCTACGAA 22901 TTCAGCTGGG TTGAGGATUT GGGUTAAATG AACCAAACCT CCCTATACAT 20951 GAAGGATACA CAGAGATGET GACAGAGAGT GGTCACTTCC GTGAGTGGAT 23001 CTCAATCAAG TOCTCTGAAG CTAAATTCAA TTTTTTTTCT TTACTAAAAT 23051 GATAAAAGTT GTTATTGGGG CYTTTGCTTG TTTATTTCGT ATAACTTAGG 23101 GCTCAGATTT TCAATGTGTC AAATGCTGAC TCACAGCATG GTTCTCCTGA 23151 CAGTTTATTT CATTTAAGGA ACTCTTCACC AGTAAGTTTA TTTACTTGCC 23201 TTGATATCTC CACACATTAA TAATAAAACT AACAAAACCT AATCTGAATT 23251 AAAATCTATC AGCTTTAGGC ATTATTTTGT GTTCTCCTTC TTTCAACATG 23301 GTAACTGGGC TCTCTTCTT AGGAGCTTGA GAAGATATGA CTGGGGTTTG 23351 TTTTTCTCTA CTTCATTTAT TATCTTTCTT TTTTCCAATC AGGTTAGTTT 23401 TTTCCTTTTT AGTAAAAGGT GCATAGTAAC TGCTTGTAGT ATTTGTTGAA 23451 CAAGTGAATA AATGAAATGA ATTAACCTAC TOTTTTCACT ACCACCCCAA 23501 CATTTCTTTC TCTCTTACTA GTGGGTGGGG TATCAGTTAT GGAATGGCAC 23551 CTCCTTCCAG AGGACTGATC ATGTCATTTT CAGCTTATGC TTCCCTTTAT 23601 GCAGTAAAGT TTCCATATTT CCATAAAGAA CAAGAAACCA AATAATCCTA 23651 ATGGATATAT AATGAACACA CAGATGAAAA TTTCACCTGC CATGCCTTTG 23701 AAAAAAGATC CCTAGCTACT TGTATTTCAT CTTATAATTA AAATCAGTCT 23751 TTTCACTTAT GTTTTCTTCA GATCTCCTGT TTTGAAGTGT ATATAGATAT 23801 CAACATAGAA ATGCAGCGTA TATTGCTATC AACTGCAGTG GAGCAGTGAT 23851 TCGTAGGTTT TCCAACATCC TTGCCTTAAG CAAACCTGCA AAATCAAAGT 23901 GTGAGCTACG TCTAAACAAT GGGAGAGGCT TTTTTTTTT TTTTAAGAGT 23951 TAGAACTAAG ACTCTCACTT CCTCCTGTGC CTCCACATTT TTGACCTTCA 24001 CATTGGGCCC CTGCATCAGA ATACAGCACC CCCTAACAGG CTCCTGTTCA 24051 GGACTCTTC TCTGGAAATA ACAGATGTTG TCTCTAGAGC TGCATAGAAC 24101 CTTAATGGAA TCATTGTGGG TCAGAGGCCC TGGATGGTGC TGGGGACCTC 24151 CCTGACCCAC AGCATCTGAC CCACATTTCC AGGTTCCTAG CGACTTGTGT 24201 CAGTAAAGAA AAAGGCACAT AGCTAAGTGG AAGAGCAGAT GAGGCTTGGT 24251 GGGAATCAGC CAGTGGTCTG CCCTAGCAAA GGTAAACAGA ACTGCTGGGG 24301 GCTTTTGGTC CTAGGCTCAC TACTCAGGGA GGCACTTTAA CATGGAATGA 24351 CCAGCAAGTT TCCTTGCTGA TCTTTTCCAC CACCACCACA AGCCTAGTAC 24401 CTCCCTCCCT CTTTGCTCTG TTGCTCTCTT CGGGAATGCA CTGGAAACCA 24451 CCTTCAGTTC TGTTTGGAAT TTTCCTATTC CTTATTCAGA AAGAGGAAGA 24501 AGCTTTTGCA TTTACTCCAA CCGTTCTACC TATTATTCCC ATAAACTTTC 24551 TGTGATCTCA TATCATTAGG CCAAATGTTA ATCTTTCTGG GAGCCAGGAG 24601 ACTGCTTTCA CATTCAGAGG CCCTGGACAT ATAGGACTGC CTCTAACTCA 24651 CTCTAACTCA GCTTATTGAC TTGAATGCAC CTTTTTAACA AGTGACTAAA 24701 AAACAAACTG TGACTATTCT CTGAAAATGA GCCTATATCT CATACTTATT 24751 TATTCTGTTT AACACTGTGA AACAAATTAA GTCCTCTGGC ACTATGTATA 24801 TACCATAAAA AGCTTATTTG TAAGCCTACT AATTGGACCA GTTTTGACAA 34851 TATTGAATAA GCACTAATTG CAGATCATAA TGTAGAATTA TAGGCTGCTG 24901 AGGAAAACAA TATCACACCA TTTGCTTTCC TCAGTTTCCT TTTCAGAATG 24951 AGTTTCATAA TGTTCACTAA TCCAATTTTT AAAATCCTTT ACAAAGTTAT 25051 TGCCTTTTCA GCCTAAACAG ATGGCCTTAA TTTTTGGTGG AGTGGTATGA 05101 AAGGAATGTC ACATGAGAAA CTGCAAGCTA TTTAGCTTGA ATTTTTTGTC 25151 ATTCATACAT GTTTCAAAAT ATATTTTACA TTTTCTCTCT TTTAAATGAG

25201 TTCCCATCTC TGCACCTTAA GTGACTTCAG AACTAAAATT TTAAAGTGAA 25251 CATCAATCAC AGCATTTOCA AAAATGTGAA CTCCTAGCTT AACCGAAGTA 25301 TTCACTTATT GGAAAGCTGA TAGAGTAATT CCACTAAGTC CAAAAAGTGT US 351 CCTCTAAAAG ATTCCCAAAGA TAAGAGTGTT TTCAACTTTG TCAAGCTGTA .5101 CAAACACAAA TGTCACTOOC TOCOTOTGCC CACAGGGATC TOTATOCAGT US451 TACAGCAGOG TAACTIGAGO AGOTGOTGOA AACTGAGGOT OTOTIGAGOO .:5301 TTCGCCTACT TATTTCAGCT GCTAAAATAG GGCTGAAATC TGTCAAGGAT :5551 CCTGAAGGGA AGGATAAGAT CCCTACTATT CAATTTAATT TAAGCTTTTA 15601 TTCAGTGCCT GCTGTGTGCA CAACACTAAG CTAGAAAGTC TGAGGAATGT 18651 TTAGATTATT AGGTOSTGTT COTTGGCTTT CATAGATTTA CAATCTATTG US/01 ATAGGGAGAG CTAAAAAGGA GAGAAAGAGG AAGGAAGGAAA MATAAAAAAGG 25751 TCAAAATTTT AAAATAGCAT TTTAAAATTT TATTTTAAAA TUTTAAAATAC USB01 CATGCAAAAT TAAGGAAAAC STAGATTCAT AAAAATTCCT TICACAATCT ASTROPORTO AATTOAGTGO TEGCOTETAA EGECETCATOO AATOEGATGA 35901 GACATGITTI GIGATCAACA AGGGITTIAC TATGITTCIT AATTATGIGT 15951 CTTGCCTGTT ATCTCTTTCT GACCGAGATT ATTTTTAGGA ATAAATTCTG 06001 AAAACTAAGA AAGTGAAAGO ATAAAATATT GTOTTATAAA ATAOGOOAAG 06051 GAAAAAATGA CACTOCATTT CAAATATCAA AAGTTAGCAT CAAGACTGCA U6101 CAAGATGAAT GTACAGTGAT GTGTTGCTTA CAAATGTGGA GATATTGTGA 06151 GAAATGCATC TTTAGGCAAT TTTGTCATTG TGCAAACACO ATAGATTGTA 06201 CTTGCAGCCT AATTGGTGGA GCCTACTATA CACTAAGGGT ATATGGCATA 16251 GCCTAGTACT CCTAGGCTAC AAACCTGTAC AGCATGTTAC TGTACTGAAT 06301 AGTGGAGGTA CCTGTAACAT AATGGTAAGT ATTTGTGTUT COAAACGTAG ..6351 AAAAGCTACT GTAAAAATAC AGTATTACAA CCTTAGGGTA TCACTGTCTT 26401 ATATGTGGTC TGTTGTTGAC CGAAATGACT ATGCTTAATA CCACTGAACT 26451 GTACACTTAA AAATGGTTAA GATGGTAAAT TCTATGTTAT GTATGTTTTA 26501 TAATAATAAA AAAATTGAAA AAAGCATCAA CATCTTTTT GGGAAAAAAG U6551 AAAAAGAAAG AAAATGCATT AGAGTGATGA GAATATTTGA AGTAATAGAT 26651 TTCGTACATC AGTGATCACA CTGTTATTTC CCAAACGACC CTTCAGCTGG 26701 ATACGACATT TCCTGATTGC AGCTGTGCTT ATTGCACTTA ATAATGTTCT 26751 GGTCATCCTT GTGCCAACTA TAAAATACAT CTTCGGATTC ATAGGTGAGT 26801 TTCAGAAAGG CTTCAATTTG GTCAACCCAA ACTCACGCCT CATTAAATGA 26851 TGGACAGGGA ACCAGTGCTG GGTCATCCAG ATCCCCGTTC TTTCTCAGGC 26901 TCATGGATTC CCTTTATCCC TGCGAGGCTC TGGTGATTGA GCTGCTCACT 26951 GTCTCTTCCT CCTAACTGAC ACTGGGAGCC ACCTTATAGG TCATTTAGTC 27001 AAGCTGCTTT TTCTGATAGA TGAGGAAACT GACCCCTATA AAAGTCAAGT 27051 CATATACCTT GGTGTGGACC CAGGATTTGG ACTTAGGTAT TAGCTCCACC 27101 ATCAGGAAAA GAGGAAGATA GATTTTACCT GCCAGAAGCT CTCTGATACT 27151 ACGAGTATCA GCTGAACATT GAAAGGTATC TTCAGAGGAA TAGGAGGTTG 27201 ATTATATAAA GTGTATTATT AGTATTTCCC CATAACTGCA TGGTCTATTA 27251 ATTTTCATTC TACTCATTGA GGGTTTACTT AAACTTTAAA CACAATCTAA 27301 AACTTTAAAA GAACCATGGG TAGGTCACTT GCAAAGTAAG AGGTGGATAG 27351 GGTGTGTCAT GAGTTCAGCC ACCTTAGTAT GTATTTATAT TACTAATCCC 27401 CTGTAAATTT GTGTTAAATT CAGCCTTTTG TTGCTTATTA TATGTTGCAT 27451 ATACTTATGC AGCTTTGATG TTAGGTACAT TTTAATTGTC TCTATAAACA 27501 TATCTTCTAT GAATAAATAA CCAAGATGAG CTTATGTGAC TTAAGTGTGT 27551 GTTTTTAGTG CTAAGTATAG GATAGCTTTA TATTTGGTTT ATTTAAAGTG 27601 TGTGCTGGCA TCTCCTTTGC TAGGAACTGC TGGGTAAGAC AFTGACCTTG 27651 CCCTGTGTTT GTCTTCTCAG GGGCTTCTTC TGCCACTATG CTGATTTTTA 27701 TTCTTCCAGC AGTTTTTTAT CTTAAACTTG TCAAGAAAGA AACTTTTAGG 27751 TCACCCCAAA AGGTCGGGGT AAGTAAACCT TGCAATTTUC CCCATTATTA 27801 GTTGTTCTTC CAACTACTTA GAATAAACTA GAAAATACAC ATAGTTCAGA 27851 AAAATGAATC AATGTACAAG AACCAAAAAT CAAAAATGGG CTAGAACTTT 0.7901 CTGGTAGCAG AGAAAGGGGA CATATTTCTG AAACTCAAAT GATTCTACTT 27951 CAAATATCAA ATATCCTGTG TTGAGTCTGT CATACATGTC AAATAGTAGT 28001 AGCCTTTCCC ACAGACACAT ATGCTTCAGG CAAATAGCAG TSTCCAATAC 28051 CAAGCTGCTG TTGTGCTATC CGTGGAAAAT CATGCAAGAA GGAATTAGGC 18101 TCCCTAGCGG TGTTATGGAA TAATTTAAAT ATTTTGGT A TGGTTGTTAG 28151 GTTTGCAAAG CCAAAGGAAA GATGTTGCTT TTGTTTTCCC TTCCATAGTA 28201 CCTGTTGTCC CTGGTGTGGA CTAAGATCCA GAACAGAACC ATTCATCGTT 18251 CTGTTAACCT CTTTAGATAC AAAATACAGT CTTATTAAAT TAGAGAGTAC 28301 ATATTTCTTT TCCATAAGAC TACTATAGAA ACAAATGCTA GAAATAATTG

28351 TTTTTCCAAT AAGGAAATAT TATCTTTCAC TOOTTAATAA AGTCATGTTA 20401 AGGCTTGAAA AGAATATTTC TTACTGAATT ACTCTGAATT TITACCITGA UH451 AGTCATITAC CTTTGGGATG TTCTGGGGAC TTCAGGATAA TITGGTATCA 24501 AAAGGTCCAC CCAGCAGCTT GCTCCCAAAT TTTAACTCTA TCTAGTCCGT 23551 OTTGOTIGGA TITTTACAGO AGTGTGACCI TGGCAAAUTA CITGICCIGI 28601 TTGTGACCTA TTTTCAGTTT GACCAATTGT GAAATGACTA CAACTATCTC .:4651 CTAGACCEAT TCTAGTGAAA AATGTTTAGT TGCTGCT.TC TPARATGTAG JETOT GATTAGGAGG TITAAGTATG IGATAAAATI TAAGGCCCCI ICI MITGITA 20751 AAATGOTGAA GTATTTTATA TGTAGGTATG FACATATATC CETATATATG UBB01 TGTGTGTATA TTATATGTAT GCACACACAC ACACACAIAT AFACACTTT UNBST TGTTGCAACA TCTATTAAGC TTTTTGGTTTT GTTTGGTTTA TAAAATTAGA 18901 ATCATATCAT ATATGCTATT CTTTTTTAAC CTGCTCT.TT TCACCTAAAA .59001 FAATAATCAC AGGGTATTCC ATCATCTTGG TGTACTA/AT CAATTAACTA 29051 TTACTCCATT GTTGAACCTG TAGGTTGTAT CTCTCCACTG TATTCCTCTT .09101 CTTTCTTCAA CTAGGATTCT AAATTGACTG ATAGGTDAGG CUTGGGCATC UMISI TGAGATATTA AGAATAATAT GGCTCAATAT ATAGATGAGA TIGGCCATATT UP201 ATGTAAACAA CTAAAAAACA AATTGTACTA AGTATGBTTT OFGTUCTCCT 29.51 AACAGAGTOT CTOTGAATTA CAGGOTTTAA TTTTCOTTGT GETTEGAATA 39301 TTCTTCATGA TTGGAAGCAT GGCACTCATT ATAATTGACT GGATTTATGA 29351 TCCTCCAAAT TCCAAGCATC ACTAACACAA GGAAAAATAC TTTUTTTTTC 19401 TATTGGAAAT GGTTACAAGT TATACTCCAA AAGATATTTG AATTATCTTG 29451 ATTGGAATGT TATTCATAGG AAATAACAGG AAGATTCCAA AGACCTTTAC 19501 CAGTAATATO ACCAGGOACO TGCAGAAGAG GAAAATUAUT GTTTTTGTCA 29551 AGGATGGTTG TGTATGTGTT TAAAATAAAA CCTGTGGTGC ACATTTCTAC 29601 CCAGGTTTTG CTAGAGCAGT GTGAGATGAT GAAGGTTTAT TTTTGCTGCT 39651 TTACGAGCAG AATAAGGGTA ACTGCATGTA ACAATCATCA GAFAGTACTC 29701 TTTCCCCTGC CGTCTCCTCA TCCTGCACCC CCTAAAAAAG TAOCAAACAT 19751 TTCCATTCTC ACAACATCAA ACAAAAATGC CCTOGT.GGA AAGCTATCAC 29801 CATTTAATGT CTTCTCTCAG TCTTGCACCA AAGTCTCTGG TCTGTTTACT 29851 AACAGAGGCA AAAGGCATGT CTTAGGAACT GTTTCTGTTT CTGTAAGGTA 29901 CATGAATGGT CAAACACCAG TCTAGAGCAT CTTATTGTTA ACAGCAAAAT 29951 AATATTTTGC CCACCCTGTT TGTGACATTG AGTTGTGACT TCTATATTCA 30001 ATAGATTTTT GTAAATGTTA AAACATCTAT ATTTAAATGT TAAAACACTA 30051 AATATAGAGA GGGGCTTTAT TTCAATCATA GAGCAACAAC AAAAATAATG 30101 CTTATAGCTA AACTGCCTGT TCTAGAAAGC ATCTGCTTTT TCATGTTATT 30151 CCTAAATCCT CTTGTCATAC TTTTGTCATT GAACAATGCT CTCCCTCTCG 30201 TCTTCCATCC TCATTCAGAA TTTTTAGAAG ACCACAATCG TGGAGATACA 30251 CTACCCAGTA TTGTTTGATA CATTTTTATT TGATAAACAT TCAGTGCAGG 30301 AAACTGTGAT TTGCTATATG TTTATGTATA TAATCTTATT CTGTAGTCAT 30351 CAGAATGTTA ATGTAAGGTA CATTTGATTT TTATTTTTTA CATGTGTAGT 30401 TTTCTTTCTT CACAGTCAAA GCATTTATAT TATTGGGGGT GGGGGCAGGG 30451 AATTAAGTTG GTGGGCTCGA AAATCCATTC ATATGTATUT GTCTACAAAT 30501 GTCTGGGGAT AATTTAAATT TGAAACCTAA GTTATATATA GTTTGGCAAT 30551 GCTCTTCTTC AATATTTACA ATAATAGGAT GATCTACAAG AAAATAAGTT 30601 TCTTTTTGCA AATTTTTATC ATACTAAAGT TGTTCTTT"A ATTTAGCATA 30651 TCTAAAATAG GAATTAGTTC AGTTTAGCTC ACACAGGTGT TTGCTGACAT 30701 TCATTGGCCA TTTAATACAG TGTTGAGTGG TTCTCCCTAA AAGTATAAGT 30751 GCTAACACTA CGAAGAAATG CACACGATCA TTOTTGGTGA CTTCTATAAC 30801 AAACTTACAT AAAATGGATT TAAAAATTCC TACTCACAGG CTAAAACTTC 30851 TGGAGTTCAC TACCTTTTTT TCAAATCATA GTAAGATCAC TTGTGTATTT 30901 TATATTTTAG TAAAGCCAAT TATGAAGTAC AAGTATCATA CACGTACTTT 30951 TGAGCTACTA TTATTTGAAA AAAATCTGCC AAATAGCATC TTTAGGATAT 31001 ATTTACATTT TCACTCATCT AAAAAGTATA CAAAAATAAA AAGTGGAAAA 31051 AGGTATCTTC TGAATGTTCA AGAGCATCCT ATAGTGCCAA ATAATAAAGC 31101 ACCATTTTTT TCTTCATAAC CAGGATTAAA ATTCATATAT ACTGCAGGGC 31151 AGACATACAT ATGATAGCTT GTGCTGATTA ATTTAAGCCC ATTTGTAAAC 31201 AGATGAAAAT TTTATTTTCT TATTTCATTT ATAAGAIGGC TCAATGTATT 31251 GGGAGGCTTC TTTTTTATTA CAGAAAGTGT ATATTGGTAT ATAATAAATG 31301 AACTTTTCAA ATGACTATGA TGTGATTTTT GATCTATTGT TAAAGAATGT 31351 TGTGTTATTT GTCCATGAAA CAAAATTTAA AATCCAAATA CTGTCTTTCT 31401 TATATTGGTT TATGTTCCAT TTTCATTGTT ACCTTTGACA CATAACTAAC 31451 ATCTATAGCC ATCATCCTGA AAATAATTGC CATCTTATTT TGGCAAAATA

```
31501 GATATTTAAT COTAAATTAT TATGATGATT ATAATTTTGG CATCACATAT
31551 ATACCACCTA GAATGAATGT GGAAGAAATG ASTCTTTTAT GGTTAGTTTG
31601 AAAGAATCCA FTGAAGATAG AAAATGAGAG AATAGAAGAA ACCTGAGAAT
31651 AGTAAAATAA AGAGCAGAGA AAATATGGGG GCAGGGAAAA CATGTGAGTG
31/01 CTAAGGATTG ATTATGAATG AACGATTAGG GGGATTGATG GATCACAGGG
31/51 TAAGTATATG CTIAACTTTA TAAGAAACTT CCACATAGTT TTCCACAGIG
31301 TTTCTACCAT STTCATTTCC ACCCGTACTA CCTACAACTT CCACTGACTC
31851 CACAGOCCTG OCAACATTTG GTGTTGTCTT TTGCATTTTA GCCTTTCTAG
31901 TGGGTCTGAA ATGGTAACTC ATTGTGATTT TCATTTCTGC TTCTGTGACA
31951 ACTAATGTTG AAAACTTTTC AAGTGTTTAA TGGTCACTCA TATATCTTCT
32001 TTTGTGAAGT GTGTATTCAA ATCTTTTGCC CATTTTTAAA ATTTAGGTTA
32051 TGTGTTTTTA TTGGGTATTT GTAGAAGCTC TTTAAATATG GATCCATGTC
3.:101 CAGATTGCCA ATATATTTTC CCAGTCTATG GTATGGTTGC TTATTTTCCT
30151 AAAGGTGTCT TAATTACATC TTTCTGGGGC CAGGTCACCA TAGCTCAAAG
32.001 TTTTGCAATT TATGTCTTAA TGAGATAATA TTAATCAGAG TGGTATAGTC
30251 AAAATTAAAT OTTTTGATGT CCTGGGCCCA TATAGGTAGG ACTGGATCAT
3.301 CTAACCAAGA TGCAAAAAAA AAAAAACAAA AAAACAAAAA TAGTACTTGG
30051 AAAAACTTAT TTTAAATTAA ACA
```

## FEATURES: Start:

3000

Exon: 3000-311% Intron: 3119-7453 Exen: 7453-7543 Intron: 7544-8039 Exon: 8040-8155 Intron: 8156-10894 Exon: 10895-10968 Intron: 10000 11107 Exon: 11438-11530 Intron: 11531-16047 Exon: 16048-16129 Intron: 16130-16215 Exon: 16216-16298 Intron: 16299-16408 Exon: 16409-16467 Intron: 16468-17301 Exon: 17302-17577 Intron: 17578-17709 Exon: 17710-17789 Intron: 17790-19073 Exon: 19074-19174 Intron: 19175-20904 Exon: 20905-21029 Intron: 21030-26649 Exon: 26650-26794 Intron: 26795-27670 Exon: 27671-27768 27769-29273 Intron: Exon: 19274-29372 Stop: 29373

## CHROMOSOME MAP POSITION:

Chromosome 12

## ALLELIC VARIANTS (SNPs):

	WITHIT	(SMES).				
DNA				Protein		
Position	Major	Minor	Domain	Position	Major	Minor
1386	T	C	Beyond ORF(5')			1.11101
2594	T	C	Beyond ORF(5')			
2757	13	Т	Beyond ORF(5')			
6107	Ç	'Γ	Intron			
6392	T	C	Intron			

FIGURE 3, page 11 of 23

						_	
9484	C	G	Intron				
10330	А	G	Intron				
10.297	G	Ä	Intron				
10331	ıĠ	A	Intron				
10536	Т	Ĉ					
11549	T	Č	Intron				
11917	:3	T	Intron				
1.840			Intron				
1./944	Ā	-	Intron				
12847	T		Intron				
13019	i O	-	Intron				
13022		-	Intron				
	ÿ	G	Intron				
13285	(3	Α.	Intron				
14461	ij	C	Intron				
15464	-	G	Intron				
15469	-	A	Intron				
15545	ή,	2	Intron				
16199	T	13	Intron				
16798	'n	0	Intron				
18103	C	1	Intron				
18421	Α	43	Intron				
18508	G	А	Intron				
18722	Ľ.	Ç.	Intron				
18775	C	G	Intron				
18951	T	Ç	Intron				
18974	T	G	Intron				
19540	Α	C	Intron				
19841	G	А	Intron				
3Ú14U	V.	r	Int.on				
20343	Ţ	C	Intron				
20519	(;	А	Intron				
20963	T	C	Exon		411	P	
21840	G	ů.	Intron		411	Г	P
22783	C	Ţ	Intron				
22787	G	Ä	Intron				
22825	T	C	Intron				
22967	A	Ť	Intron				
23248	Α	Ğ	Intron				
23764	G	T	Intron				
23765	C	T	Intron				
24432	A	Ğ					
24538	Ċ	G	Intron				
24693	T	C	Intron				
24819	Ĉ	T	Intron				
25743	C	T	Intron				
26044	G	C	Intron				
26555	G		Intron				
27886	A	A	Intron				
31884	A T	C	Intron				
32229	T	C	Beyond (				
32223	1	А	Beyond (	DRF(3')			
Context:							
OULLEAL:							

Context:

DNA

Position

1386

 $\tt CTTATTTGTATTTTTTTTTTTTAATCTCTTCATGCTATAATTTGAGTGATTTCCACAGA\\ TCTGTCTTTCAATTTTATAAGTCTTCCTTCAGCTGAGTTTTTTTAAATTTCAATGATTCT\\$ 

FIGURE 3, page 12 of 23

ATACATAGTCGAGCATTTTATATAAAAACAACTAAAAAGTCTGTGACATTTTGCAGTATA
AAAATGCAATGGCAGCAGCAGGCCTTATTAATTGAGCCTCTTGGAAATGTGGCTGGTCCT
AGGTCCGTAGCCTCAAAGGCCCTGGCTTGTAACTGCAGGAGCTGACCAGCACAGCTCTAT
AACCAAGTTGTACATCTTCTAGCCTGTGTCCAAGAAAACCAGAATCACAACGCTCTGTGG
ATAGTGACATCTTAAAGTTTTCTTTCCCTCCCAACTCTTTTGCCAGTTCATTGAATTGCT

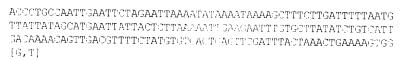
GCAACATTTATATCACAAATATGTGCTGTTTATGTTCTGAATATCACATATGATTAGTAA
TCACACAGCTATTTGAGGGCTAAGCATCAGGACTATAAATATTTGTATTGTGTTAGTGCT
TTGATTGAACTCTTTTATGTATAATATTCTTCAGCTGAATGGGTTTTTATATCAACTTTA
CTTTTATATAAGCCATGTTTTGAAATAAACTAGGATTTTAATAATCTGAATTTTAATAGC
TATGTATGTAGTCATATATTTGTATGCTTTTGTAATGTGCTTACCTCTAAGACAAAAAAA
[C,G]

CTGCCTTTCCTTATTAATTATACCATACCATTAAAATGAATTAGGAAGTTACAGATCACTGAGAGATGAATAGAAATAGGAAAAACTTCCCCCCAATCCCACAGATCATAGATCATCTTCATGAGAGAAGAATGTTCCACTTTTAAAATGAGGGCCTCATTTTAGGCTTATAAACACTTAGCAGATGAATTTGGTCAGAACAATTAAATCACTAAACATCATGGGGTGTTTTTGTGTGTCTAAGTAGCCCAGAGACCAGTATTAAAGCACTTAATTTATAGCAAGTGACACAGTATTTTAAAG

FIGURE 3, page 13 of 23

10280	ATAAGAGTGCAACATAGCTACAGGGGTTATAAAATTTATAATTCATGGTCCAAATGTACA TTTCTA STAFTGATTTCATTGGGAATTACCAA SGGATTAGATCAATTGTGGGGAAA STGT ATTTTTTAAAAATAAACAAAGATAAAGATFTTTTTTTTTT
10297	CTACAGGGGTTATAAAATTTATAATTCATGGTCCAAATGTACATTTGTAGTATTGATTTC ATTGGGAATTACCAAGGGATTAGATCAATTGTGGGAAAGGGAAGTGTATTTTTTAAAAATAAACAAAGGAAAAGAAAG
10331	AAATGTACATTTGTAGTATTGATTTCATTGTGAATTACCAAGGGATTAGATCAATTGTGGGGAAAGTGTATTTTTTAAAAATAAACAAAGATAAAGATTTTTTTT
10536	TACCAATCCGTATAAACATAAAGAAACCATACATGCATGTGGGGATCCTAACACCAGAAA TGACTCTGAATGCAAAAAAAAAA
11548	ACCATTGATTCTTGTTTGGAGAACATTTTGATATATTGCTTATTGGTTTTTGAGGTTGCA TCTTTTGGGCTTATAATTTCTATATGATGTTTTATTACATGTTTTGAGACTCCAGCATGGA ATTATATGACAAAAATATTTTAGTCATTAAAACAATCTCTTTAACAAGGCTATTTTATCT TTGATTGTAGGGTCTTTGATTTATGAAAAATTAGGAGAAAAAGGCATTTGGATGGCCGGGA AAAATTGGAGCTTTTGTTTCCATTACAATGCAGAACATTGGAGGTAAGGGGATATACTTT [T, C] CAATGGATCCCATAAACTTTCTATAGCGTSTTCAATAAATAAGAAAACTTATGGCAATAA ACAGGCACTTTAGATACAGAAAAATTGCTACTTATAGTTCTTAAAATTTAAAATGATAGT TTCTTAAATAGGTTTGTGTCCTGCTTTAATTAAAAACAGCAATATCTAAGAATGAAATAA CATATAAAACCCTGCCAATTGAATTCTAGAATTAAAAATTAAAATTAAAAGCTTTCTTGAT TTTTAATGTTATTATAGCATGAATTATTTTTAAAAATTGAAGAATTTGTGCTTTATAT
11917	TTTAGATACAGAAAAATTGCTACTTATAGTTCTTAAATTTTAAAATGATAGTTTCTTAAA

 ${\tt TAGGTTTGTGTCCTGCTTTAATTAAAAACAGCAATATCTAAGAATGAAATAACATATAAA}$ 



[A,G]
CACACACACACACACACACACACACACACACACACACA
GETCTATTGCTCCTCTAGACTGGTAAATACTAGTTCCTCTGGGCTCTCATGGTCCTGTTT
GCATCTAGTATGTTACTGTTTCTAAAGGATATTTCAAAAACACTTGAGTAGAGAAATAAGC
TE PTGGAGTCTGATGGACCTGAATTTGAGT CTGPPP CTGTCAGTATCTGTGAACTTGGGA
AGATCACTGTACTCCTTTGTCTGATTTTTTCATGTATAAAAATTACCTTACAAAGGCTAT

ATTTTTTCATGTATAAAAATTACCTTACAAAGGCTATTGTGAGGATGAAATAAGGTAACA
TATGGCACATAATAAGTGTTCTGTATATGCTTCTCCCCCCGGTTCTCTGCTTCCATA
TCCATGTCTCTGGAGTTGCCTGAATTATTTTTTAAATAGGCATTTAAAAAATTATAAAAAC
AAATATATGATGATTGTGAAAAACTAAAATAAATATAAAATTAAAATTACCAAGAAAA
CTTTATGTCAGTCATCCTCAGAAATAACTACTCATAGGTTTTCCCCCTATGCCTAATTCAA

TATOGAGCATTTCATAGGATTGCCTTATAGTTGGTCTAATTTAACAACTGAAATAACCAG
GCATAAGCATAATTAACCCTGGACTCAAGAAGTTGAGTGGCAGCACCTCAGCTGTGGTCC
AAAGCATAGCCACTACTACTGCTTCTAAACAATGGAATAAAGTATAAAGCGTCTCTCAGT
CAAGCCTCACACAGGTAAGAGGCGTGACCTTAAAGGAGTAAGATGAAATATCGTAACATC
ACCCCAGAAATAATGCTCTCACTTTGGTTACTTTATTTGATTAGTTGATATTTGGCATAA
[G, C]

AGAAATCACTTGTATTTCTCTATTTAACAACTCTACATTTAGAACACTTAATTTTCTCAA TCCCCTAAAAAATTAACATTTACTGCAGATGTTTTCACATTAACAGATTAATGTCTGGAT CATTCTGAATTTTTGAAGACCAAACATGTTAACATCACTGACATCACTGAAAACCAGCAA TTAATAGCTCTAACATTGAATGGTACCTCACCAAGGCCAGCTTAACCAAAATAATG

TGAGTTCTATTTTTAACTGAATCTTTTGGCCATGTGTCAACAAATTAACGTTATCCTTCA
CCAAATGGGTGGGCTTGAAAAAGGCGTGATGCATAAATATTTACAGTTGTAGGCAAAATT
GTAATGTTATGTATATGAATACATATTCATTTTTTCAGGGAGAAGGCTTGTAGATTTCAT
CAAGAAATCTTTCACAAGAGATAGATAATCATTCATGTATCACTTACCTAGATGCTCATGA
AATTTTGCCACTTTATATAATTCCTTAGTTAGCCAAAAGGAGAGTAAGATGAAGAGGGGG
[-,G]

AAAAAAAAACTTCTTTGACAAAGATGGAGAGAAGCTGTCATCTCTTGTATTCTTTTATC
AATCCAGGAAGCCTTTGGTTTTGACAATAAGTGGTCTGAGACTTTGTGACTCCTCAGAT
AGGTCCCGGAGGACTAGATTGGTGCCCATCTGCAGAAAACCAGAGGGGATATATTGACTC
TGCAGATCTGCCCTTTGATTCTGCCATCTCTCAGCTGGCCCATGCCTTTTGTTGCCAGAC
TACTGCCCAAGTTATAGACACTAACACAGGCACACTGAGTATGGGCTATGTTGATTTATA

TOTATTTTAACTGAATCTTTTGGCCATGTGTCAACAAATTAACGTTATCCTTCACCAAA
TGGGTGGGCTTGAAAAAGGCGTGATGCATAAATATTTACAGTTGTAGGCAAAATTGTAAT
GTTATGTATATGAATACATATTCATTTTTTCAGGGAGAAGGCTTGTAGATTTCATCAAGA
AATCTTTCACAAGAGTAGATAATCATTCATGTATCACTTACCTAGATGCTCATGAAATTT
TGGCACTTTATATAATTCCTTAGTTAGCCAAAAGGAGAGTAAGATGAAGAGGGGGGAAAA
[-, A]

AAAAACTTCTTTGACAAAGATGGAGAGAAGCTGTCATCTCTTGTATTCTTTTATCAATCC AGGAAGCCTTTGGTTTTGACAATAGGTCTCGGAGAGACCTTTGGTTTTGGCTCAGATAGGTCCCGGAGGACTAGATTGGTCCCATCTGCAGAAAACCAGAGGGGATATATTGACTCTGCAGATCTGCCCTTTGATTCTGCCAGCTCTCAGCTGGCCCATGCCTTTTGTTGCCAGACTACTGCCAAGTTATAGACACACAGGCACACTGAGTATGGGCTATGTTGATTTATAACTAA

> TGACAATAAGTGGTCTGAGACTTTGTGTACTCCTCAGATAGGTCCCGGAGGACTAGATTG GTGCCCATCTGCAGAAAACCAGAGGGGATATATTGACTCTGCAGATCTGCCCTTTGATTC

> > FIGURE 3, page 16 of 23

18421

TOCCATCTCTCAGCTGGCCCATGCCTTTTGTTGCCAGACTACTGCCCAAGTTATAGACAC TAACACAGGCACACTGAGTATGGGCTATGTTGATTTATAACTAATGAGGGCACCACCCTA CAACTGCAGCTTCACTGTAAACTTTGGAGCAGGATTTAACACAGAAATCAGCCCCGATACT

FOR TOTAL TOTAL TARGET AND THE ARTIGORACT AND TOTAL TO

> O ATATTOACATGAATCAGCCATTCTCTCTTTACACATTCCACCATTAAGAGA GACAAG AA GAGTGGGATTAAAGAAGAAATCCTCCTCTCTCTAGGCCCCTGACAAAAGAGGGAATTTCT TGCACTATCATGAATGCCAAAATTTATAAAGCATTTCCCCAAAGAGGTAAAGGAGAAGGA AAAAAAGTTTTGAAGACCCATGTCACCTTAGTTTGAAGAAATAAGGAAATGATCATCTTT OTTATGGAAGGGCATGAAAGAGGGTGCGAAGGATTCTTGCAAAATATTGTCUTGTAACT

TATTATATAATTATTCTCAACTTCTTTAATCTAGACATCAGTAACCTCAGTCTTCACTCTCTACTAATAATAGCAAAACTTTCCCCATAAATTCTGATTTACCTCATAAAAAATTTCAGAACACTTTCAAGAACTTTCAGAACTTTCAGAAAATTTCAGAAATTTCAGAAACTTTCAGAAACTTTTCAGAAACTTTTCACAAACCTTTTCACAAACCTTTTCACAAACCTTTTCCACAACCCTTTTCAGAATGTCTCGACTTTTCACAACCCTTTTCAGAATGTTCTCAGAACCTTTTCACAACCCTTTTCACAACCCTTTTCACAACCCTTTTCACAACCCTTTTCACAACCCTTTTCACAACCCTTTTCACAACCCTTCTCACAT

ACAAGAGATGTGTCAATATCTGTCCTTTGTTGTCTTATTTGTACAATAGAGTCACTGGGT
AGAAAATCTTGTTTCCTCCAGCTGATGGTCTATGGTTCATTTGTATTCTTTTCCTTTTGA
AGTTGTTGATATTTGCTTGGGAACAAAGGATATGAACTCATTATAGCTGTTTTCCTCTTT
CCTTTAAGGGAGGATATTATATAATAATTCTCAACTTCTTTAATCTAGACATCAGTAACC
TCAGTCTTCATTCTCACTAAATAGCAAAACTTTCCCCATAAATTCTGATTTACCTCATAA

AAATTTCAGAACACTTTCAAGTATTTGATGTCTTTGATTTACTTTGAAAATTACAFGTAGGTATCAGAACACTTCAAGAACCTGTTTGATACTTTGACACCAGGTTCCTTCTAGAACTGTTTCAGAAGCTTTCCAGAAGCTCTGGACTCTTGGCAGACTCTTTGCTGACTCTTAGGTTTTTTCCTCATTTAAAGTCATCTCATTATGAAATGCAAAAGCTTTCTTATGTTAGGACCCTGTTTCATCTTATGTTAATTATTCTTATTCAGTGGGCAAGCTTACTGACCTACGGGCACCTTACTGACCTACGGTG

FIGURE 3, page 17 of 23

18722	TATTATATAATAATTCTCAACTTCTTTAATCTAGACATCAGTAACTCTCAGTCTTCATTCTCACTAAATAGCAAAAACTTTCCCCATAAATCTGATTTACCTCATAAAAAAATTTCACAACACTTTCAAGTAITTTGATGTCTTGATTTACTTTGAAAATTTACTTCACAACACTTTCAAGTAITTTGATGTCTTGGAAGATTACTCTTGAAAATTTACATGTTACACAAGCTTTCCAGAAGCTTTCCAGAATGATTTTCCCTCAGAAGCTTTCTAGAATGGTTCTCAGAAGCTTTTCCCTCATTCCCTCATTCCCTCATTCCTTCAGACTCTCATTCCTCATTCCTCATTCAT
18775	TUATTOTOACTAAATAGCAAAAOTTTOCOCCATAAATTCTGATTTACOTOACAAAAATTT CAGAACAOTTTCAAGTATTTGATGTTTGATAAAAATTATCAGAACAATTTCAAGAACAATTTTGATGTTTGATAAAAAATTATCAGGAACCAGTTTCAAGAACAATTTTTGATGAAAAATTTTTGAGAAAAATTTTTTTAAAGTCATGACTTTTTGAGAATGGTTTTTTAAAAAAAA
18951	CAGAAGCTTTTCAGGTAGTCTGGACTCCTGGCAGTAGTACTTTGCTGACTCTACTAGGTT CTTTTCCTCATTTAAAGTCATCTCATTATGAAATGCAAAAGCTTTCTATCTTACGTAGCCT GTTTCATCTTATGTTAATTATATTCTTATTCAGTGGGCAAGCTTACTGACCTACGTGAA ATACACTCTTCCTCTTCTAGGGAAATGATTGTTTTTAAGACTGAAGACTAAAGTTTA AAAAATGGAAATGATCCTCATTAGCCTCTCTAAGACAAATTTAAATCAGCTATAAGTTTA [T,C] GTACTAAATATGTCTTCATGATTAGCAATATAGATATACTTTTTTATTATTATTTTCCATT TTGAAAAGTGATTTTTTTTTT
18974	ACTCCTGGCAGTAGTACTTTGCTGACTCTACTAGGTTCTTTTCCTCATTTAAAGTCATCT CATTATGAAATGCAAAAGCTTTCTATGTTAGGAGCCTGTTTCATCTTTATGTTAATTATA TTCTTATTCAGTGGGCAAGCTTACTGACCTACGTGAAATAGACTGTTCCTCTTCTAGGCA AATGATTGTTTTTAAGACTGAAGGACTAGTGTTTAAGAAAAATGGAAATGAATCCTCATT AGCTCTCTAAGACAAATTTAAATCAGCTATAAGTTTATGTACTAAATATGTCTTCATGAT [T, G] AGCAATATAGATATACTTTTTTATTATTTTTCATTTTGAAAAGTGATTTTTTTT
19540	GGTATGATTTTAAAAATTGCCTTCTATAAGCATGCTCTATAGATGACACATATTCAATT AATATACTATTTTAGTTTTGTCACTTGACCTGAGGAAATGGGGCCTGATTCAGCCTGGCT AACAAGTTACAAGAATTTGTGAATTAACACCTATTTTATAAAAAATATCCCCAAAACAAA ATTATTTTCCTCTAGGGATAGATGATATTTCTCTGGCTAGACTCCATAGTCCAACTCAGG CTACAAGTGATGAGAATCCACTTGCATGTGATAAAGCTCCTTTGATGGAATTATTA [A, C] CTGCCACACAAATAGCAGGGAAACTGCCAGGTCCTCAAGTTTGAATTTGCCTCCTTTA CCAGTCAAGTCA

# FIGURE 3, page 18 of 23

 $\tt CTCTCTGGCATCTTAACAGTTCACAAAGGGAGTAGGATTGTACTCCTTCCATGAAGTGTG$ 

 $\tt CTGCCACACAAATAGCAGGGAAACTGCCAGGTCCTCAAGTTTGAATTTGCCTCCTTTTACCAGTCAAATCTGGGAGCTTGGGACTTTAGGTAAAATTTCTGACATATCCCATTCCTATTCCTATTCCTATTCCCATTCCTATTCCCATTCCTATTCCCATTCCTATTCCCATTCCTATTCCCATTCCTATTCCTATTCCCATTCCTATTCCTATTCCCATTCCTATTCCTATTCCTATTCCCATTCCTATTCCCATTCCTATTCTATTCCTATTCCTATT$ 

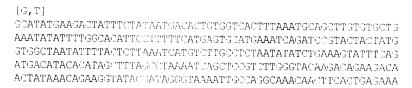
TATTTGTTATACTAAATGATTTCCTAAGAAAGAGGACATGACAGAATTTCCTTCAATCT AAGAATGCACCCACAAAAAAAAACTGACCACAAAAAGGCACATTAGATTATGCCTCCCAACAACTTTCCCTCTCGCAACAACTCTCACAAAAGGGAGTAGGATTGTACTCCTCCATGAAGTTTG

OCACATAAACAGATTTCA ESGAA ECACATATTGACCTGGTAGCATATETTTACAT SAATC AGTGTATCAATATAAATATACTTTTTGTATAAACCTCCTTTTAAAGTT PTTAACTTAATTTTTTTTTTTTATATTGACAAATTGACAAAATTGACAAATTGACAAATTGACAAATTGACAAATTGTGAGGAAAA SATTC AGGAGTAGGCCACCATTTSCTTAG STTTTTTTTTTTTTTTGACTAATATTTGACTAATATTTGACTACTAATAACCAAACATGTGACGA

20170 TATTGACCTGGTAGCATATOTT FACATGAATCAGTGTATCAATATAATATATTTTTTGTA
TAAACCTCCTTTTAAAGTTFTTAACTTAATTTTTTCTTACTGACTTGGTAAATTGAATT
GCATGTATGACAAATTGTGGGGGAAAAGATTCAGGAGTAGGCCACCATTTGCTTAGGTTT
TTTTTCTATTTGACTAATATTTGACTATTAACCAAACATGTGCTTTAGATTGGCCATTAA
CTTTTTGCCGGTTGTGAAATAATGAATGACGAGGTCAATACTACTGAAGGTATTTCACT
[A, C]

20343 TAGGTTTTTTTCTATTTGACTAATATTTGACTAACCAAACATGTGCTTTAGACTGG
GCATTAACTTTTTGCCGGTTGTGAAATAATGAATGACGAGGTCAATACTACTGAAGGTAT
TTTCACTACTTTTTGTCTGATCTTGAGGTGAAAATCCAACTACGCTTGATTCCATAGATA
TTTTCTTGTTATTTGTGTTGSAGTCCTGAATGAAGGTGTTTTCAAGTAGGGCTGCATCT
TCGTCTTAGAGTAGTACCCACTGGGAGACCATCTAAAAATTATACTAATTTATCCCTGCA

20519 GATATTTCTTGTTATTTGTGCTTGGAGTCCTGAATGAAGGTGTTTTCAAGTAGGGCTGC
ATCTTCGTCTTAGAGTAGCCACTGGGAGACCATCTAAAAATTATACTAATTTATCCC
TGCACGTTACTTATACTTATTTTAATGAGTTTCATAAGACAAGCAAAAACTTGAAAGAGC
CCAAAAATATCTGTTTTAGTGTGTGATGGAGTCATAGTTGTTGAGCTTGAAAAATTGGT
AGCAATCATTCATCCTAGAGTTTACACACTGGGTTTGTAACCTGCATCAGGAGTGGCTGC
[G, A]

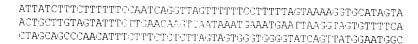


TGAGAAATAAAGCACTGATA FALATCT GA DUA FOAGGAACAGCAATAGTGTGAAACATT AGATGCCATTAGAACATTGACCA FAA GAACCAGAGTTCAGAAAATGACTGACTGC TG FCCTTCATTATGTATTT CUACTGAACA FTAGCATTATGAAACATTTTGCACATTATC CTGTCCTCACCCTTGCAATH FLACATT FA FATAATCTGTGTAAGTGCTCCACTGC DCCAC AGAGTCATAAGTCCCTGGGACTTGGT GATGTGCACAGTGACTGCCCCACGCCTGCACTGC CCAC [C,T]

22787 AAATAAAGCACTGATATAAATCT GACCATCAGGAACAGCAATAGTGTGTAAACATTAGAT
GCCATTAGAACCAAAATTGACCATAAGAACCAGAGTTCAGAAAAATTGACTAACTGTGTC
CTTCATTATGTATTTCCACTCAACATTAGCATTTATGAAACATTTTGCACATTATCCTGT
CCTCACCCTTGCAATGTTACATTTATATATCTGTGTAAGTGCTCCACTGCCCCACAGAG
TCATAAGTCCCTGGGACTTGGTGATGTGCACAGTGCCCACAGAGGGTGAGCTCTGTC
[G, A]

CAATAGTGTGTAAACATTAGATGCCATTAGAACCAAAATTGACCATAAGAACCAGAGTTC
AGAAAAATGACTAACTGCTGTCCTTCATTATGTATTTCCACTCAACATTAGCATTTATGA
AACATTTTGCACATTATCCTGTCCTCACCCTGCAATGTTACATTTATATAATCTGTGTA
AGTGCTCCACTGCCCCACAGAGTCATAAGTCCCTGGGACTTGGTGATGTCACAGTGACT
GGCACAGAGGGTGAGCTCTGTCGTGCTTGGGAAGAAAAATGGTCTTCAAATGAATCTTGC
[T,C]

> > FIGURE 3, page 20 of 23



GGAATGCACTGGAAACCACCTTCACTTCTGTTTGGAATTTTCCTATTCCTTATTCAGAAA GAGGAAGAAGCTTTTGCATTTACTCCAACCGTTCTACCTATTATTCCCATAAACTTTCTG TGATCTCATATCATTAGGCCAAATGTTAATCTTCTGGGAGCCAGGAGACTGCTTTCACACTCCAGAGGCCCTGGACATATAGGACTGCCTCTAACTCACCTCTAACTCAGCTTATTGACTT GAATGCACCTTTTTAACAAGTGACTAAAAAAAACAAACTGTGACTATTCTCTGAAAATGAGC

CCTAGTACCTCCCTCCTCTTTGCTCTGTTGCTCTCTCGGGAATGCACTGGAAACCACC
TTCAGTTCTGTTTGGAATTTTCCTATTCCTATTCAGAAAGAGGAAGAAGCTTTTGCATT
TACTCCAACCGTTCTACCTATTATTCCCATAAACTTTCTGTGATCTCATATCATTAGGCC
AAATGTTAATCTTTCTGGGAGCCAGGAGACTGCTTTCACATTCAGAGGCCCTGGACATAT
AGGACTGCCTCTAACTCACTCTAACTCAGCTTATTGACTTGAATGCACCTTTTTAACAAG
[T,C]

FIGURE 3, page 21 of 23

24819	AA COGTTCTACCTATTATTCCCCATAAA CTTTCTSTSATCTCATATCATTACGCCAAATGT TAATCTTTCTGGGAGCCAGSAGACTGCTTTCACATTCAGAGGCCCTGGACA FATAGGACT GCCTCTAACTCACTCCAGCCTTATTCACTCAGACTGCACCTTTTCTCTTTAACAAACTGACTTATTCTTCTTT TAACACAAACTGTGACTATTCTCTCSAAAATGAGCCFATAFCFCATTTATTCTTTTTTTTTTAACACTGTGAAACAAACTTAACTCCTCGGCACTATGTAFATACCATAAAAAGCTTATT [C,T] GTAAGCCTACTAATTGGACCAGTTTTGACAATATCAAACACCTTATTTCCAGATCATA ATGTAGAATTATAGGCTGCTGAGGAAAACAATATCACACCATTTCTTTC
25743	TATCCAGTTACAGCAGCGTAACTTGAGCAGCTGCTGCAAACTGAGGUT CTCTTGACCCTT CGCCTACTTATTTCAGCTGCTAAAATAGGGCTGAAATCTGTCAAGGATCCTGAAGGGAAG GATAAGATTCCTACTATTCAATTTAATTT
26044	AAAAACGTCAAAATTTTAAAATACCATTTTAAAATTTTATTTTAAAATGTTAAATACCAT GCAAAATTAAGGAAAACCTAGATTCATAAAAAATTCCTTTCACAATCTTGTGTAAATCCAT TCAGTGCTTGCCCTTAATGTCTCATCAGGTCTGATGAGACATGTTTTTTTT
26555	AGTACTCCTAGGCTACAAACCTGTACAGCATGTTACTGTACTGAATAGTGGAGGTACCTG TAACATAATGGTAAGTATTTGTGTCTCCAAACGTAGAAAAGCTACTGTAAAAATACAGTA TTACAACCTTAGGGTATCACTGTCTTATATGTGGTCTGTTGTTGACCGAAATGACTATGC TTAATACCACTGAACTGTACACTTAAAAATGGTTAAGATGGTAAATTCTATGTTATGTAT GTTTTATAATAATAAAAAAAATTGAAAAAAGCATCAACATCTTTTCTGGGAAAAAAGAAAA [G, A] GAAAGAAAATGCATTAGAGTGATGAGAATATTTGAAGTAATAAAGTCAAAAACAAA GAAATGATCTTGCCTTTGAACTTTCTTGTTTAAGATTCGTACATCAGTGATCACACTGTT ATTTCCCAAACGACCCTTCAGCTGGATACGACATTTCCTGATTGCAGCTGTGCTTATTGC ACTTAATAATGTTCTGGTCATCCTTGTCCAACCTAAAAATACATCTTCGGATTCATAGG TGAGTTTCAGAAAGGCCTTCAATTTGGTCAACCCAACC
27886	GGTTTATTTAAAGTGTGCTGGCATCTCCTTTGCTAGGAACTGCTGGGTAAGACATTGA CCTTGCCCTGTGTTTGTCTTCTCAGGGGCTTCTTCTGCCACTATGCTGATTTTTATTCTT CCAGCAGTTTTTTATCTTAAACTTGTCAAGAAAGAAACTTTTAGGTCACCCCAAAAGGTC GGGGTAAGTAAACCTTGCAATTTCCCCCATTATTAGTTGTTCTTCCAACTACTTAGAATA AACTAGAAAATACACATAGTTCAGAAAAATGAATCAAATGTACAAGAACCAAAAATCAAAA [A, C] TGGGCTAGAACTTTCTGGTAGCAGAGAAAAGGGGGACATATTTCTGAAACTCAAATGATTCT ACTTCAAATATCAAATATCCTGTGTTGAGTCTGTCATACATGTCAAATAGTAGTAGCCTT TCCCACAGACACTATGCTTCAGGCAAATAGCAGTGTCCAATACCAAGCTGTTGTGC TATCCGTGGAAAATCATGCTAGGAGAAGGAATTAGGCTCCCTAGGGTGTTTATGGAATAATTT AAATATTTTGGTCATGGTTGTTTAGGTTTGCAAAGCCAAAGGAAAGATGTTGCTTTTTTTT
31884	CTTTTATGGTTAGTTTGAAAGAATCCATTGAAGATAGAAAATGAGAGAATAGAAGAAACC TGAGAATAGTAAAATAAAGAGCAGAGAAAATATGGGGGGCAGGGAAAAATATGTGACTGCTA

TGAGAATAGTAAAATAAAGAGCAGAGAAAATATGGGGGCAGGUAAAACATGTGAGTGCTA

